

EXHIBIT A

**Expert Report Concerning A Study of Possible Adverse Impact of the 1999 and 2002
Examinations for Firefighter, NYC Fire Department**

Joel P. Wiesen, Ph.D.

November 23, 2007 (errata corrected as of 1/4/08)

Table of Contents

Section One: Background Information	Page 4
A. The Topics of This Report	Page 4
B. Summary of My Opinions	Page 4
C. Documents and Computer Files Reviewed	Page 4
D. My Professional Background and Some Related Information	Page 6
E. An Overview of the City's Hiring Process for Firefighters	Page 8
F. Evaluation of Adverse Impact	Page 9
Section Two: Adverse Impact Analyses: 1999 Exam	Page 13
1. Adverse Impact in Appointments by Race: Exam 7029 (1999)	Page 13
2. Adverse Impact in Appointments of Males by Race: Exam 7029 (1999)	Page 16
3. Adverse Impact in Passing the Written Test: Exam 7029 (1999)	Page 18
4. Adverse Impact in Passing the Physical Performance Test: Exam 7029 (1999)	Page 19
5. Adverse Impact in Males Passing the Physical Performance Test: Exam 7029 (1999)	Page 21
6. Relationship Between Written Test Score and Taking the PPT: Exam 7029 (1999) .	Page 22
7. Adverse Impact in Passing Overall Exam: Exam 7029 (1999)	Page 24
8. Average on Written Test and PPT: Exam 7029 (1999)	Page 25
9. Adverse Impact in Rank Order Placement On the List: Exam 7029 (1999)	Page 27
10. Adverse Impact in Appointment Among Those Certified: Exam 7029 (1999)	Page 30

11. Adverse Impact in Appointment Date: Exam 7029 (1999)	Page 31
12. Adverse Impact in CNS Disposition Action Codes: Exam 7029 (1999)	Page 35
13. City's Evaluation of the Adverse Impact in Passing the Written Test: 1999 Exam ..	Page 36
Section Three: Adverse Impact Analyses: 2002 Exam	Page 38
14. Adverse Impact in Appointments by Race: Exam 2043 (2002)	Page 38
15. Adverse Impact in Appointments of Males by Race: Exam 2043 (2002)	Page 40
16. Adverse Impact in Passing the Written Test: Exam 2043 (2002)	Page 41
17. Adverse Impact in Passing the Physical Performance Test: Exam 2043 (2002)	Page 43
18. Adverse Impact in Males Passing the Physical Performance Test: Exam 2043 (2002)	Page 45
19. Relationship Between Written Test Score and Taking the PPT: Exam 2043 (2002) .	Page 46
20. Adverse Impact in Passing Overall Exam: Exam 2043 (2002)	Page 48
21. Average on Written Test and PPT: Exam 2043 (2002)	Page 49
22. Adverse Impact in Rank Order Placement On the List: Exam 2043 (2002)	Page 51
23. Adverse Impact in Appointment Among Those Certified: Exam 2043 (2002)	Page 54
24. Adverse Impact in Appointment Date: Exam 2043 (2002)	Page 56
25. An Evaluation of the Effect of Typos and Similar Data Errors	Page 58
Section Four: Census Comparisons	Page 60
26. Adverse Impact of One Year of College by Race: Census 2000	Page 60
27. Adverse Impact in Male Applications: 1999 Exam	Page 61
28. Adverse Impact in Male Applications: 2002 Exam	Page 63
Section Five: Opinions Summary	Page 65
29. Opinion Summary	Page 65

Attachment A: Documents Reviewed from Vulcan Society et al. v City of NY et al. . . .	Page 67
Attachment B: Resume of Joel P. Wiesen, Ph.D.	Page 70
Attachment C: List of Papers and Publications	Page 71
Attachment D: Replication of the Basic Adverse Impact Analysis: Exam 7029 (1999) . .	Page 74
Attachment E: Replication of the Basic Adverse Impact Analysis: Exam 2043 (2002) . .	Page 75

Expert Report Concerning A Study of Possible Adverse Impact of the 1999 and 2002 Examinations for Firefighter, NYC Fire Department

Joel P. Wiesen, Ph.D.
November 23, 2007

My name is Joel Wiesen. I have been retained as an expert by plaintiffs-interveners in *Vulcan Society et al. v City of New York et al.* This report was prepared at the request of counsel.

Section One: Background Information

A. The Topics of This Report

In this report I describe the results of adverse impact analyses I undertook of two open competitive examinations for Firefighter in the NYC Fire Department: Exam 7029 (the "1999" exam) and Exam 2043 (the "2002" exam). These analyses consider the adverse impact of various aspects of these exams including, for example, pass-fail decisions, rank ordering of applicants, and date of actual hiring.

B. Summary of My Opinions

A brief summary of my opinions is provided here concerning my independent review of data and related written material I received. I opine separately on the 1999 and the 2002 exams.

i. The 1999 Exam Has Large and Very Highly Statistically Significant Adverse Impact on Black Applicants

Exam 7029 has practically and very highly statistically significant adverse impact against black applicants in appointments. Several other aspects of the test (e.g., written test score, Physical Performance Test pass-fail, overall exam pass-fail, time to appointment) show statistically significant average black-white differences, to the detriment of black applicants.

ii. The 2002 Exam Has Large and Very Highly Statistically Significant Adverse Impact on Black Applicants

Exam 2043 has practically and very highly statistically significant adverse impact against black applicants. Several other aspects of the test (e.g., written test score, overall exam pass-fail, time to appointment) show statistically significant average black-white differences, to the detriment of black applicants.

C. Documents and Computer Files Reviewed

In preparing this report I reviewed over 26 documents provided to me by counsel. Some of these materials were specifically requested by me, as I saw a need for them to allow me to

more fully evaluate the impact of the examinations.¹ A list of the written documents I reviewed is provided in Attachment A.

I. Computer Data Files Reviewed

I reviewed and analyzed data in four Excel files, two for each Exam. These computer files were given to me by counsel. I understand that the City provided these files. The file names and a brief description of their contents follows.

a. Applicant Data, 1999 Exam

I understand that this Excel file, named "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS", contains data on the 17,212 persons who took Exam 7029 (the "1999 exam") open competitive written test for Firefighter, 1,749 of whom were identified as black and 12,916 of whom were identified as white, for a total of 14,665 applicants who were identified as either black or white. The data in this file includes but is not limited to: applicant identification information, ethnicity/race, gender, test date, written test raw score, date of physical test, written test standardized score, Physical Performance Test (PPT) raw score, PPT standardized score, combined weighted standardized (z) score, transformed score (FA)², veterans credit, legacy credit, residency credit, adjusted final average, list number, first certification date, first disposition action, last certification date, and last disposition action. Applicants who were appointed as Firefighter from Exam 7029 can be identified by the code "APP" in either the first or last disposition action field. In analyzing this and the other three data files, I considered only data for applicants who are identified as either black or white, ignoring all other applicants.

b. Appointment Data, 1999 Exam

I understand that this Excel file, named "CID_DOJ_7029_qualapptd.xls", contains data on 3,226 persons who were appointed from the 7029 (1999) written test. The information provided in this Excel file includes but is not limited to: applicant identification, gender, and appointment date.

c. Applicant Data, 2002 Exam

I understand that this Excel file, named "2043 Spreadsheet FINAL.xls", contains data on the 17,825 persons who took Exam 2043 (the "2002 Exam") open competitive written test for Firefighter, 1,393 of whom were identified as black and 13,878 of whom were identified as white, for a total of 15,271 applicants who were

¹I have noticed, and I have been informed by counsel, that certain information has not yet been provided by the City. If germane additional data is received, I intend to incorporate it in a supplemental report. I note this more specifically later in this report.

²This is the wording in the heading of a data column in the City's Excel file. I did not receive a key to such names, abbreviations or acronyms from the City.

identified as either black or white. This data file includes the same data fields as the corresponding file for the 1999 exam, as described above.

d. Appointment Data, 2002 Exam

I understand that this Excel file, named "CID_DOJ_2043_qualapptd.xls", contains data on 2,131 persons who were appointed from Exam 2043 (the 2002) written test. This data file includes the same data fields as the corresponding file for the 1999 exam, as described above.

ii. Integrity of the Data Files

I accept the data files listed above as accurate because, although they were given to me by counsel, I understand that they originated with the City and so reflect official City records. Nevertheless, I found a small number of apparent errors and inconsistencies in the files. For example, for each exam there is one applicant who is listed as taking the exam twice (on the same day) and achieving the same score. (I do not include the duplicate applicants in the analyses reported below.) As another example, the number of people listed as appointed as Firefighter from Exam 7029 (1999) is not the same in the applicant and appointment data files for that exam (2,902 versus 2,921 appointments respectively, for black and white groups combined). As a result, in the analyses below the number of applicants appointed will depend on the which data file is being analyzed, a decision that is driven by the contents of the data files (e.g., appointment date only appears in one file, certification information only appears in another file). There may be reasonable reasons for all or some of the apparent errors and inconsistencies in the data files. In any case, these apparent errors and inconsistencies involve a relatively very small number of applicants, and so do not substantially affect the results of the statistical analyses I report below.

D. My Professional Background and Some Related Information

This section summarizes my professional background in general and as it relates to this legal proceeding.

I. Education, Professional Experience and Qualifications

I was awarded a Ph.D. in Psychology from Lehigh University in 1975. My major field of doctoral study was psychology and my minor field of study was psychometrics. Since 1975, I have worked in that area of psychology known as industrial psychology, particularly in the development of fair, valid personnel assessment methods and instruments for the assessment, selection and promotion of employees. I have taught college courses in industrial psychology, experimental methods, and statistics at the undergraduate and graduate levels. For over 15 years I worked for the Department of Personnel Administration, later called the Human Resources Division, of the Commonwealth of Massachusetts. I was hired to lead the effort to validate all civil service examinations and, for some of my tenure, I was in charge of all test development and validation for civil service jobs in the Commonwealth. Also, while employed there, I was charged with developing two job performance evaluation

systems: one for all state managers (covering about 3,000 employees) and the other for all non-managerial state employees (covering perhaps 50,000 employees). Since I left the employ of the Commonwealth of Massachusetts, my professional work has consisted mainly of consulting to large organizations on personnel assessment and selection matters, and developing new written tests for government, business/industry and education. In the course of my professional career I have conducted several thousand adverse impact analyses in employment settings: hiring, promotion, and layoff. I have served as President of the International Personnel Management Association Assessment Council (IPMAAC), an organization of over 500 testing professionals, and as President of the New England Society for Applied Psychology (NESAP), a Boston-area association of applied psychologists. From 1989 to 2006 I served on the Program Committee for at least eight annual conferences of the Society for Industrial and Organizational Psychology, a national, professional group of several thousand members. I have served as an expert for both plaintiffs and defense in employment discrimination cases, including race, age, and sex discrimination cases, a few of which have gone to trial. I have served as an expert consultant on testing matters for the Massachusetts Attorney General's Office, the Connecticut Department of Administrative Services, the Connecticut Attorney General's Office, and the US Department of Justice, as well as several large private sector organizations. I have lectured and published in the areas of personnel assessment, employment discrimination, and civil service examining. I have made presentations before professional groups, including several invited addresses. I am licensed as a psychologist in Massachusetts and Pennsylvania. A copy of my resume is attached to this report. (See Attachment B.)

ii. Publications and Papers

In the past 10 years I have authored publications and presented papers at professional meetings, sometimes co-authored with others. A list of these 35 publications and papers is presented in Attachment C.

iii. Compensation

I am being compensated at my usual rate of \$225 per hour for all work on this matter in calendar year 2007 (schedule to increase to \$250 per hour for work performed in calendar year 2008).

iv. Court Appearances and Depositions

Within the past four years I have appeared as an expert at trial or by deposition in connection with the following matters.

Terry Urban vs. Haliburton Energy Services, Inc., February, 2004
Charleston, WV
Testified for the claimant at a binding arbitration hearing.

Jacob Bradley, et al. v. City of Lynn, et al., May, 2006

Federal District Court, Boston, MA
Testified for plaintiff at trial.

Bridges v The Alcohol Beverage Control Commission
Massachusetts Division of Administrative Law Appeals, November, 2006
Testified for defense at a hearing.

Boston Police Superior Officers Federation v. Boston Police Department and the
Massachusetts Human Resource Division, June, 2007
Boston, MA
Testified for the appellant before the Massachusetts Civil Service Commission in
January, 2007 and June, 2007 and was deposed in February, 2006.

E. An Overview of the City's Hiring Process for Firefighters³

The process the City used to hire Firefighters from the 1999 and 2002 entry-level exams began, in some respects, with the announcement of the examination via a written "Notice of Examination" that states such information as: the job title, duties and salary; the application date, fee, and how to apply; the education and experience requirements and other minimum requirements to take the examination or to be appointed as Firefighter (e.g., a driver's license); and a description of the written test and the physical test. Applicants take the written test and, if they pass they are notified and are scheduled to take the physical ability test. The test scores for applicants who pass both tests are combined using a statistical method designed to give equal weight to the written and physical tests and to yield a nominal 0 to 100 range of scores. Then various special point credits are added to the scores of individuals, such as five points for NYC residents, if applicable. Based on the "Adjusted Final Average" the names of the passing applicants are placed on a rank order list called the "Eligible List." The order on the list is determined by the applicant's adjusted final average, except that two people who receive exactly the same final examination rating are placed on the list in numerical order of the last five and then the first four digits of their social security numbers. The City may take six or more months to establish the Eligible List, and the list generally has a life of four years. Upon establishment of the list, candidates are notified of their exam scores and rank on the list. The Fire Department (FD) uses the Eligible List create a list of certified candidates to invite to continue the screening process. Invitations are issued following the rank order of the Eligible List. The FD screening process includes: checking to see if the individual meets the minimum requirements as stated on the examination announcement; reviewing documents provided by the applicants, including references. In addition, it is my understanding that a Personnel Review Board (PRB) was asked to review the files of some certified applicants and, depending on the action of the PRB, some certified applicants were not hired, and these applicants are identified in the data file by the disposition action code CNS (considered not selected) and are not appointed from

³This description is based on my review of the disposition action codes of the data files, as well as documents listed in Attachment A and others found on the City's website.

that certification . The City's rule of one-in-three allows the FD to appoint any of the top three ranked individuals on the Eligible List to an available position (i.e., job).

F. Evaluation of Adverse Impact

I clarify here the definition of adverse impact and the use of statistical tests to evaluate adverse impact.

I. Definition of Adverse Impact

Whenever many people compete for relatively few jobs, some will be hired and others not. I conducted adverse impact analyses to answer this question: Is there statistically reliable indication that black applicants were adversely affected by the screening and selection criteria related to the 1999 and 2002 Firefighter exams.

Adverse impact is defined in the federal *Uniform Guidelines on Employee Selection Procedures (UGESP, 1978)*, in Section 3.4, paragraph D, which begins as follows:

"Adverse impact and the "four-fifths rule." A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact. Smaller differences in selection rate may nevertheless constitute adverse impact, where they are significant in both statistical and practical terms or where a user's actions have discouraged applicants disproportionately on grounds of race, sex, or ethnic group."

The prototype adverse impact or disparate impact analysis compares the "bottom line," tallies and rates of hire for black and white applicants. (The terms adverse impact and disparate impact are used synonymously here and in general.) To conduct an adverse impact analysis, one must identify the relevant group of candidates. This identification should be based on the nature of the challenge. In the current matter, the challenge is to the screening and selection criteria used in appointing Firefighters based on the 1999 and 2002 civil service examinations, with a claim of employment discrimination against black applicants. So I examined whether the screening and selection criteria used to appoint Firefighters based on the 1999 and 2002 entry-level examinations have adverse impact on black applicants. I have looked at the bottom line hiring for blacks and whites, but also gone beyond that global view and separately evaluated the impact of various other aspects of these two examinations and the subsequent screening process, such as passing rates and placement on the eligible lists, as explained below. The adverse impact analyses for the two examinations are reported separately, starting with the 1999 exam.

ii. Understanding the Statistics of Adverse Impact Analyses

Adverse impact analyses use statistics in two very different ways: to describe data and to make inferences based on data. We all have encountered and have some familiarity

using statistics to describe numbers (as with averages, percents, or just tallies). Using statistics to make inferences in a formal way is not a tool the general public is familiar with.⁴ This short introduction will explain the essence of inferential statistics.

a. Using Statistics to Understand Variability in Data

The actions or events we observe in the real world have causes that we know or posit, but often the causes are not clear and it appears that chance looms large, for example: will a fish take your bait, will you catch a cold, will a flipped coin come up heads, will you get a job, will a certain candidate get elected? We can count the number of times we catch a fish, catch a cold, get heads, or get a job, or poll more people about the candidate. But when should we be sure that we have identified a good fishing spot, an unhealthy lifestyle, a fair coin, or a winning candidate? How many observations do we need to provide a specified level of certainty? By quantifying the level of uncertainty, inferential statistical methods help us make decisions based on limited data.

b. The Law of Large Numbers

We all have an intuitive understanding that more data lets us make more certain judgements: we fish on more days or we flip the coin more often or we poll more people. We understand intuitively that the variability in the data will “average out” if we have more data. But sometimes we must deal with the data we have and not collect more. Inferential statistics is based in large part on quantifying the variability in a limited number of observations, and making numeric probability statements related to the trends we may see in our data.

c. A Coin Tossing Example: In Theory

Let’s say we will observe exactly 100 tosses of a coin and will have to decide if we think the coin is fair based on the observations. How many heads will it take before we think the coin is biased? If we knew how much variability in a fair coin, we would be better able to evaluate our observed 100 tosses. If it is common to get more than 60 heads on tossing a coin 100 times, then if we get 60 heads or less we might most reasonably conclude that we have no reason to say the coin is biased in favor of heads. Inferential statistics provides a mathematical basis for describing the expected amount of variability in observations.

d. A Coin Tossing Example: In Practice

Say you tossed a coin 1,000 times and got 550 heads and 450 tails. Without inferential statistics you may be unsure what to think about the fairness of the coin. Inferential statistics can help by informing you that a fair coin would give you an outcome as extreme as 550 heads only about 1 time out of 500. With that

⁴The early roots of modern methods of statistical inference go back several hundred years, to a time when statisticians were called on to evaluate the fairness of games of chance.

information most people would conclude that the data offer strong evidence that the coin is biased. This example lends itself to use of a statistical test known as the Chi Square test (which I described below), and I used a Chi Square test to arrive at the probability of 1 in 500.

e. Why Not Just Look at the Percent or Number of Heads?

If you toss a coin twice and get two heads, you have gotten heads 100% of the time. Although that is as far from 50% as you can get, the probability is 1 in 4 of getting two heads on tossing a fair coin twice; something that is not unusual. In our 1,000 toss example, 550 heads is 55% out of 1,000, a difference on only 5%, yet the probability of getting 550 heads in 1,000 tosses is quite low. Clearly the percent of heads by itself is not a good metric for probability. Similarly, initially we were not sure what to decide about the coin that gave us 550 heads out of 1,000, or 50 heads more than expected. But if we tossed a coin 20 times and got 20 heads we would be sure the coin is biased, even though 20 heads is only 10 more heads than we would expect by chance. So number of heads by itself is not a good metric for probability. Statistical tests calculate numeric values based on our data that are designed to allow us to answer the question, based on chance alone how unusual it is to get a result as extreme as we observed.

f. The Components of a Report of a Chi Square Test

The report of a Chi Square statistical test has at least three main components.

(1) The Chi Square value

The Chi Square test statistic is calculated based on the observed data. The Chi Square statistic can vary from zero to infinity. Other things being equal, a larger Chi Square value indicates that the outcome is less likely to have occurred solely by chance.

(2) The Associated Probability

For every Chi Square value it is possible to find the probability. I used the Chi Square value and the degrees of freedom (see below) to find the value of 1 in 500 in the coin toss example above.

(3) The Degrees of Freedom

Finding the probability for the value of Chi Square that is calculated from the observed data is computationally tedious, and depends on the number of groups being compared and the number of categories involved. (In choosing between four presidential candidates, there would be four categories. In the coin toss, there are two.) Fortunately, the hard work has been done for us. A technical term, "degrees of freedom," as used with Chi Square test, reflects the number of groups and categories, and is used by statisticians to calculate the probability from the Chi Square value. The reader can safely focus on the probability value and ignore the degrees of freedom.

g. Other Statistical Tests

The Chi Square test is widely used with categorical data. Other statistical tests are needed for other situations, such as comparing average scores of two groups. To compare the numeric averages of two groups, either the F test or the t-test is appropriate. I indicate in the tables below when I use either of these tests, as is professionally appropriate in the reporting of such analyses. Both of these tests involve interpreting a test statistic calculated from the data. Both tests use degrees of freedom to help find the probability associated with the calculated value of the test statistic. This probability is the probability that an outcome as extreme as the observed outcome would occur by chance alone.

h. What is Statistical Significance?

In the coin toss example we said that a probability of 1 in 500 (or .002) is too rare to accept as happening purely by chance, and so we decided that the coin was biased. The most usual standard of statistical proof uses .05 as the critical value. If the probability of an outcome as extreme as the observed outcome (or result) is .05 or less, we say the result is statistically significant and we say that is improbable that the outcome occurred due to chance alone. In many commonly used statistical tests, a probability of .05 corresponds to 1.96 standard deviation units. Perhaps the next most commonly used criterion is .01, which corresponds to 2.56 standard deviation units. This is the basis for the US Supreme Court's reference to "two or three standard deviations" as a general rule for evaluating statistical significance (*Castaneda v. Partida*, 430 U.S. 482, 496-497, n. 17).

I. What is the difference Between a One and Two Tailed Test?

A statistical test asks the question, what is the probability of an outcome as unusual as we have observed. Often when we have two possible outcomes we are concerned about differences in both directions (e.g., biased in favor of heads and biased in favor of tails). Sometimes when we have two possible outcomes we are interested in identifying an unusual outcome in one direction. For example, if we have an established medical procedure, we may be interested in deciding if a new procedure is better than the existing method, and we may have no interest in deciding if it is worse. When we are interested in identifying unusual events in either direction, we use a two-tailed test. When we are interested in identifying events in one direction, we use a one-tailed test. Usually experts for plaintiffs like to use one-tailed tests, since it is easier to achieve statistical significance with one tailed tests, and because the usual employment discrimination case involves adverse impact against minorities, and ignores the possibility of bias in favor of minorities. Experts for the defense usually prefer two tailed tests, for the same reason. In the analyses I conducted for this report, I use the conservative, two-tailed approach unless otherwise noted.

Section Two: Adverse Impact Analyses: 1999 Exam

1. Adverse Impact in Appointments by Race: Exam 7029 (1999)

The most basic evaluation of adverse impact involves an evaluation of the number and percent of black and white applicants appointed (or hired). I did an adverse impact analysis of appointments from the 1999 exam, as follows.

a. Number of Appointments: Exam 7029 (1999)

I tallied the number of persons appointed based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". For the sake of this analysis, and other similar analyses below, I included as appointed those applicants with first or last disposition action codes that indicated they were: appointed, appointed from a promotional list, offered a job and turned it down, or failed to report after accepting appointment (disposition action codes APP, AOL, DEA, and FRA).⁵ For the sake of this analysis, and other similar analyses below, I excluded those applicants with disposition action codes that indicate they could not be appointed for reasons that were generally of a voluntary nature. The excluded disposition codes indicated that an applicant: was deceased, failed to report for investigation, failed to report for interview, was not qualified for appointment, was overage for the position, or was underage at time of appointment (disposition action codes DCE, FRI, FTR, NQA, OVA, UNA). These exclusions were based on a combination of the first and last disposition code: the first disposition code had to be one of those just listed, and the last disposition code had to be either one of those just listed or blank (indicating the applicant was not certified again).⁶

These tallies are presented in Table 1a below, with the associated percents. This table shows 6.7% of the black applicants were appointed, as compared with 23.6% of the white applicants.

⁵The data file available to me contained only disposition action codes for the first and last certifications. Although these are arguably the most important certifications, if and when I get disposition code information for any other certifications, I may redo this analysis.

⁶I replicated this analysis counting only appointments (disposition code APP), and without the exclusions just noted, and the results were virtually unchanged as to adverse impact and statistical significance. This analysis is summarized in Attachment D.

Table 1a. Appointments from Exam 7029, by Racial Group			
Group	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
Black	115 (6.7%)	1,596 (93.3%)	1,711 (100%)
White	2,986 (23.6%)	9,688 (76.4%)	12,674 (100%)
Total	3,101	11,284	14,385

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

b. Adverse Impact Ratio for Appointments by Race: Exam 7029 (1999)

As shown in Table 1b, the adverse impact ratio for appointments is .28, which I got by dividing the percentage of blacks appointed (6.7%) by the percentage of whites appointed (23.6%). This result tells us that black applicants were appointed at a rate only 28% of the rate that white applicants were appointed, or that black applicants had roughly 1/4 the chance of being appointed as white applicants. A statistical analysis of this adverse impact ratio is reported in Table 1b.

Table 1b. Adverse Impact: Appointments from Exam 7029						
Black: Percent Appointed*	White: Percent Appointed*	Adverse Impact Ratio	Probability ⁷	Chi Square ⁸	Standard Deviation Units ⁹	Shortfall ¹⁰
6.7%	23.6%	.28	< .0001 ¹¹	251.8	15.9	254

* All applicants offered a job are considered appointed for the sake of this analysis.

c. Probability and Statistical Significance of Appointments: Exam 7029 (1999)

The adverse impact ratio of .28 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 1b, the column titled "Probability"). This probability corresponds to 15.9 standard deviation

⁷The usual way social science researchers interpret a Chi Square value is based on the probability that the result would have occurred due to chance or random causes alone. The most usual standard is: if the probability associated with the observed data is less than or equal to 0.05, the result is called statistically significant.

⁸Chi Square is the statistic that was calculated based on the data in order to determine the probability value associated with the result. Larger Chi Square values are seen when the results are less likely to have occurred by chance. The Chi Square value is reported here, as is customary in reporting statistical analyses. Here, and for all similar statistical analyses reported below, the Chi Square value reported is corrected for continuity, as is recommended practice. Here, and for all similar analyses below, the "degrees of freedom" (d.f.), a technical statistical term, is equal to 1. The d.f. is reported because that is standard practice in reporting statistical analyses, and it may be useful to statisticians reading the report. The column labeled "Probability" simply states the probability, as determined based on the Chi Square value and the d.f.

⁹Here, and for all similar statistical analyses reported below, I give the number of "standard deviation units" that correspond to the Chi Square value, as the US Supreme Court has referred to "two or three standard deviations" as a general rule for evaluating statistical significance (*Castaneda v. Partida*, 430 U.S. 482, 496-497, n. 17). The higher the number of standard deviation units the less probable the outcome is due to chance. The term "standard deviation" is used in other ways in some other contexts, but here it reflects the probability of the results occurring based on chance.

¹⁰The shortfall indicates the number of additional black applicants that would have been appointed if blacks had been appointed at the same rate as whites.

¹¹The actual probability is much smaller than .0001. To write the actual probability would involve some 50 zeros to the right of the decimal point, followed by a 1. The usual convention is not to write out such small values, but just to indicate that the value is less than some small number, such as I have done here.

units.

d. Practical Import of Appointments: Exam 7029 (1999)

The practical importance of hiring 6.7% of black applicants versus 23.6% of white applicants is obvious: there is disproportionate hiring, with a shortfall of 254 black appointments. The adverse impact ratio (.28) also fails the .80 rule of thumb of the *UGESP*. (Here and elsewhere I interpret an adverse impact ratio of less than .80 to indicate obvious practically significant adverse impact.) The City hired white applicants at about 3½ times the rate of black applicants.

e. Opinion on Appointments: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in appointments.

2. Adverse Impact in Appointments of Males by Race: Exam 7029 (1999)

Since very few female applicants were appointed, and there were proportionally more black female than white female applicants, I wanted to be sure the analysis just described was not inadvertently measuring adverse impact based on gender rather than strictly on race. Including the same disposition codes as indicating appointment, and using the same exclusions as in the analysis above, 2.1% of the female applicants were appointed while 21.7% of the male applicants were appointed; and 8.6% of the black applicants were female, while only 1.8% of the white applicants were female. Therefore I did an analysis for male applicants only. I did not do any adverse impact analyses for appointment of female applicants, because only 5 black or white females were appointed, which is a relatively small number for statistical analysis. The adverse impact analysis of appointments from the 1999 exam considering only male applicants follows.

a. Number of Males Appointed from Exam 7029 (1999)

I tallied the number of males appointed based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS" as described in the previous analysis. These tallies are presented in Table 2a below, with the associated percents.

Table 2a. Males Appointed from Exam 7029, by Racial Group			
Group	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
Black	112 (7.2%)	1,451 (92.8%)	1,563 (100%)
White	2,981 (24.0%)	9,452 (76.0%)	12,433 (100%)
Total	3,093	10,903	13,996

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

b. Adverse Impact Ratio for Males Appointed from Exam 7029 (1999)

As shown in Table 2b, the adverse impact ratio for appointments is .30, which I got by dividing the percentage of blacks appointed (7.2%) by the percentage of whites appointed (24.0%). This result tells us that black male applicants were appointed at a rate only 30% of the rate that white male applicants were appointed, or that black male applicants were appointed at less than 1/3 the rate of white male applicants. A statistical analysis of this adverse impact ratio is reported in Table 2b.

Table 2b. Adverse Impact: Males Appointed from Exam 7029 (1999)						
Black: Percent Appointed	White: Percent Appointed	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
7.2%	24.0%	.30	< .0001	226.9	15.1	233

c. Probability and Statistical Significance of Male Appointments: Exam 7029 (1999)

The adverse impact ratio of .30 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 2b, the column titled Probability). This probability corresponds to 15.1 standard deviation units.

d. Practical Import of Male Appointments: Exam 7029 (1999)

The practical importance of hiring 7.2% of black males versus 24.0% of white males is obvious: there is disproportionate hiring, with a shortfall of 233 black male appointments. The adverse impact ratio (.30) also fails the .80 rule of thumb of the UGESP. The City hired white males at about 3 1/3 times the rate of black males.

e. Opinion on Male Appointments: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in appointment of males.

3. Adverse Impact in Passing the Written Test: Exam 7029 (1999)

There are many possible reasons for the adverse impact just shown in the appointment of black firefighters, ranging from failing the examination, to low scores on the written test and PPT that comprise the examination, to low placement on the eligible list, to differential impact of the City's post-exam processing and screening, to applicants' decisions to drop out of contention. In the next several analyses, we will go through the examination process basically in chronological order of its administration and the subsequent use of the eligible list. To pass the examination an applicant must pass both the written test (with a score of 84.705) and pass the PPT. I did an adverse impact analysis of the pass-fail decisions for the written test of the 1999 exam.¹²

a. Number Passing and Failing the Written Test: Exam 7029 (1999)

I tallied the number of applicants passing and failing the written test alone based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". These tallies are presented in Table 3a below, with the associated percents.

Table 3a. Pass-Fail Written Test of Exam 7029, by Racial Group			
Group	Number Pass Written (percent)	Number Fail Written (percent)	Total (percent)
Black	1,054 (60.3%)	695 (39.7%)	1,749 (100%)
White	11,613 (89.9%)	1,302 (10.1%)	12,915 (100%)
Total	12,667	1,997	14,664

b. Adverse Impact Ratio for Passing Written Test: Exam 7029 (1999)

As shown in Table 3b, the adverse impact ratio for passing the written test is .67, which I got by dividing the percentage of blacks passing the written test (60.3%) by the percentage of whites passing (89.9%). This result tells us that black applicants passed the written test at a rate of 67% of the rate that white applicants did, or that black

¹²Sometimes the pass-fail point has little practical import, as in cases where there are so many applicants who passed that only high scoring applicants are considered for appointment. That was not the case for the 1999 exam. For the 1999 exam, 14 applicants with the lowest passing score on the written test, 84.705, were appointed.

applicants passed the written test at about 2/3 the rate of white applicants. A statistical analysis of this adverse impact ratio is reported in Table 3b.

Table 3b. Adverse Impact: Pass-Fail Written Test of Exam 7029						
Black: Pass Written	White: Pass Written	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
60.3%	89.9%	.67	< .0001	1149.1	33.9	457

c. Probability and Statistical Significance of Passing Written Test: Exam 7029 (1999)

The adverse impact ratio of .67 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 3b, the column titled Probability). This probability corresponds to 33.9 standard deviation units.

d. Practical Import of Passing Written Test: Exam 7029 (1999)

The practical importance of 60.3% of black applicants passing the written test versus 89.9% of white applicants is obvious: there is disproportionate passing of this hurdle, with a shortfall of 457 black passers. This adverse impact ratio (.67) also fails the .80 rule of thumb of the *UGESP*. The City passed white applicants on the written test at about 1½ times the rate of black applicants.

e. Opinion: Pass-Fail Written Test: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in passing the written test.

4. Adverse Impact in Passing the Physical Performance Test: Exam 7029 (1999)

In addition to passing the written test, an applicant must pass the Physical Performance Test PPT (with a score of 75) to pass the exam. I did an adverse impact analysis of the pass-fail decisions¹³ for the PPT of the 1999 exam, as follows.

a. Number Passing and Failing the PPT: Exam 7029 (1999)

I tallied the number of applicants passing and failing the PPT alone based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". Here, and in subsequent analyses of PPT scores, I assumed an applicant took the PPT if he or she had a non-zero entry in the field "Date of Physical Test" or had a raw PPT score other than zero. The tallies are presented in Table 4a

¹³Sometimes the pass-fail point has little practical import, as in cases where there are so many applicants who pass that only high scoring applicants are considered for appointment. That was not the case for the 1999 PPT: 129 applicants with the lowest passing score on the PPT were appointed.

below, with the associated percents.

Table 4a. Pass-Fail PPT of Exam 7029, by Racial Group			
Group	Number Pass PPT (percent)	Number Fail PPT (percent)	Total (percent)
Black	373 (68.2%)	174 (31.8%)	547 (100%)
White	5,289 (77.5%)	1,539 (22.1%)	6,828 (100%)
Total	5,662	1,713	7,375

b. Adverse Impact Ratio for Passing the PPT: Exam 7029 (1999)

As shown in Table 4b, the adverse impact ratio for passing the PPT is .88, which I got by dividing the percentage of blacks passing the PPT (68.2%) by the percentage of whites passing the PPT (77.5%). This result tells us that black applicants passed the PPT at a rate of 88% of the rate that white applicants did. A statistical analysis of this adverse impact ratio is reported in Table 4b.

Table 4b. Adverse Impact: Pass-Fail PPT of Exam 7029						
Black: Pass PPT	White: Pass PPT	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
68.2%	77.5%	.88	< .0001	23.9	4.9	47

c. Probability and Statistical Significance of Passing the PPT: Exam 7029 (1999)

The adverse impact ratio of .88 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 4b, the column titled Probability). This probability corresponds to 4.9 standard deviation units.

d. Practical Import of Passing PPT: Exam 7029 (1999)

The practical importance of 68.2% of black applicants passing the PPT versus 77.5% of white applicants is obvious: there is disproportionate passing with a shortfall of 47 black applicants. Although the adverse impact ratio does not fail the .80 rule of thumb of the *UGESP* the shortfall is of practical significance. Only 100 black applicants were appointed as entry level Firefighters over the entire life of the list. Therefore, a shortfall

of 47 black applicants is of practical significance.¹⁴

e. Opinion on Passing the PPT: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in passing the PPT.

5. Adverse Impact in Males Passing the Physical Performance Test: Exam 7029 (1999)

Since very few female applicants were appointed, and there were relatively more black female than white female applicants, I wanted to be sure the analysis just described was not inadvertently measuring adverse impact based on gender rather than strictly on race. Specifically, 23.0% of the females passed the PPT while 77.3% of the males passed the PPT; and 8.5% of the black applicants were female, while only 1.8% of the white applicants were female. It might be that the PPT's adverse impact on black applicants was due to the higher proportion of black female applicants than white female applicants. Therefore I did an adverse impact analysis of the PPT from the 1999 exam for male applicants only. I did not do any adverse impact analyses for passing the PPT by female applicants, because only 5 black females and 12 white females passed, which are relatively small numbers for statistical analysis. The adverse impact analysis for males on the PPT follows.

a. Number of Males Passing and Failing the PPT: Exam 7029 (1999)

I tallied the number of male applicants passing and failing the PPT alone based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". These tallies are presented in Table 5a below, with the associated percents.

Table 5a. Males Passing the PPT of Exam 7029, by Racial Group			
Group	Number Pass PPT (percent)	Number Fail PPT (percent)	Total (percent)
Black	368 (69.7%)	160 (30.3%)	528 (100%)
White	5,272 (77.9%)	1,493 (22.1%)	6,765 (100%)
Total	5,640	1,653	7,293

¹⁴Elsewhere I report 115 black applicants appointed. But there I include as appointed 15 black applicants who were offered appointment but declined (and so were not actually appointed).

b. Adverse Impact Ratio for Males Passing the PPT: Exam 7029 (1999)

As shown in Table 5b, the adverse impact ratio for males passing the PPT is .89, which I got by dividing the percentage of blacks appointed (69.7%) by the percentage of whites appointed (77.9%). This result tells us that black male applicants passed the PPT at a rate of 89% of the rate that white applicants did. A statistical analysis of this adverse impact ratio is reported in Table 5b.

Table 5b. Adverse Impact: Males Passing the PPT of Exam 7029						
Black: Pass PPT	White: Pass PPT	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
69.7%	77.9%	.89	< .0001	18.5	4.3	40

c. Results and Interpretation for Males Passing the PPT: Exam 7029 (1999)

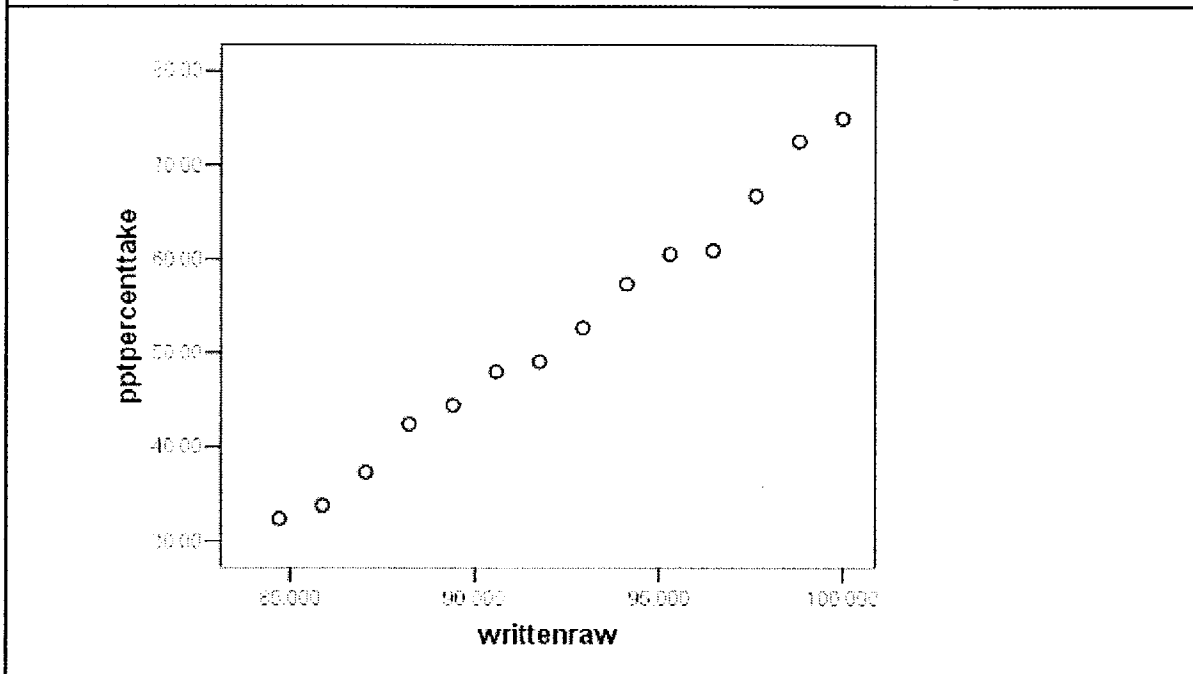
This analysis is virtually the same as the one described in Tables 4a and 4b. My comments there apply here.

6. Relationship Between Written Test Score and Taking the PPT: Exam 7029 (1999)

All applicants who passed the written test were allowed to take the PPT, but not all did so. A tally of the applicants who passed the written test shows that proportionally fewer black applicants took the PPT than white applicants (51.9% as compared to 58.8%, see Table 6). This reflects the general pattern that applicants who score lower on the written test are less likely to take the PPT. The clear correlation between score on the written test and the percent of applicants taking the PPT may be seen in Figure 1. The relationship between two such scores may be described in statistical terms by the "correlation coefficient," a statistic that can range from minus 1 to plus 1, with zero indicating no correlation and 1 indicating a perfect relationship. In the current case, the correlation is .995, which is extremely high, and also statistically significant ($p < .001$).

Table 6. Applicants Taking the PPT, by Race		
Group	Number Taking PPT (percent)	Total Number Passing the Written Test
Black	547 (51.9%)	1,054
White	6,828 (58.8%)	11,613
Total	7,375 (58.2%)	12,667

Figure 1. Relationship Between Written Test Score and Percent Taking the PPT



Notes: Pearson correlation, $r = .995$, $p < .001$.

Writtenraw = score on written test

pptpercenttake = the percent of applicants at any written test score taking the PPT

a. Practical Import of the Correlation Between Written Test Score and Taking the PPT: Exam 7029 (1999)

The relationship described above strongly suggests that the reason proportionally fewer black applicants took the PPT than white applicants is that black applicants generally received lower scores on the written test. So the (passing) written test scores affected the PPT scores in a way that goes beyond the intended examination weights. If black applicants had gotten higher scores on the written test, we would expect more black applicants to have taken the PPT and been hired.

b. Opinion on the Correlation Between the Written Test Score and Taking the PPT: Exam 7029 (1999)

There is a practically important and highly statistically significant relationship between the written test score and taking the PPT that serves to deter black applicants from taking the PPT.

7. Adverse Impact in Passing Overall Exam: Exam 7029 (1999)

To be included on the eligible list, which is required to be considered for appointment, an applicant must pass the overall exam (both the written test and the PPT).¹⁵ My adverse impact analysis of the pass-fail decisions for the 1999 exam follows.

a. Number Passing and Failing Overall Exam 7029 (1999)

I tallied the number of applicants passing and failing the exam based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". These tallies are presented in Table 7a below, with the associated percents.

Table 7a. Pass-Fail Overall Exam 7029, by Racial Group			
Group	Number Pass Exam (percent)	Number Fail Exam (percent)	Total Number of Applicants (percent)
Black	373 (21.3%)	1,376 (78.7%)	1,749 (100%)
White	5,289 (41.0%)	7,626 (59.0%)	12,915 (100%)
Total	5,662	9,002	14,664

b. Adverse Impact Ratio for Passing Overall Exam 7029 (1999)

As shown in Table 7b, the adverse impact ratio for passing the exam is .52, which I got by dividing the percentage of blacks passing (21.3%) by the percentage of whites passing (41.0%). This result tells us that black applicants passed the exam at a rate of 52% of the rate that white applicants, or that black applicants passed the exam at about 1/2 the rate of white applicants. A statistical analysis of this adverse impact ratio is reported in Table 7b.

¹⁵Sometimes the pass-fail point has little practical import, as in cases where there are so many applicants that pass that only high scoring applicants are considered for appointment. That was not the case for the 1999 exam.

Table 7b. Adverse Impact: Pass-Fail Decisions of Overall Exam 7029						
Black: Pass Exam	White: Pass Exam	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
21.3%	41.0%	.52	< .0001	249.5	15.8	302

c. Probability and Statistical Significance of Passing Overall: Exam 7029 (1999)

The adverse impact ratio of .52 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 7b, the column titled Probability). This probability corresponds to 15.8 standard deviation units.

d. Practical Import of Passing Overall: Exam 7029 (1999)

The practical importance of passing 21.3% of black applicants versus 41.0% of white applicants is obvious: there is disproportionate passing, with a shortfall of 302 black passers. This adverse impact ratio (.52) is substantially less than the .80 rule of thumb of the *UGESP*. The City passed white applicants at about 2 times the rate of black applicants.

e. Opinion on Passing Overall: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in passing the exam.

8. Average on Written Test and PPT: Exam 7029 (1999)

I report here an overview of the average written test scores and PPT scores by race.

a. Average Test Scores: Exam 7029 (1999)

I calculated the mean scores (averages) shown in Table 8 based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS".

Table 8. Mean Scores, by Race and Gender: Exam 7029 (1999)					
Test Score*	Black Mean	White Mean	Probability	F Statistic (d.f.)	Standard Deviation Units
Written Test	84.1	92.7	< .0001	1,560.0 (1, 14662)	39.5
PPT**	83.6	87.8	< .001	37.0 (1, 7373)	6.1
PPT: Males**	84.1	88.0	< .001	31.0 (1, 7291)	5.6

* A higher score is better.

** These averages were computed only for applicants who took the PPT.

b. Adverse Impact in Average Test Scores: Exam 7029 (1999)

The adverse impact ratio described in the *UGESP* is not applicable here. To measure the adverse effect, I calculated the difference between the average test scores for black and white applicants. These differences are all in favor of white applicants.

c. Probability and Statistical Significance of Average Score Differences: Exam 7029 (1999)

All the mean score differences shown in Table 8 are highly statistically significant. Differences this great would occur by chance less than one time in a thousand (and in some cases, substantially less than that - see Table 8, the column titled Probability). These probabilities correspond to from 5.6 to 39.5 standard deviation units, depending on the test score being considered.

d. Practical Import of Average Score Differences: Exam 7029 (1999)

The practical importance of black applicants, on average, scoring lower than white applicants on the scores for both parts of the exam is that, other things being equal, relatively fewer black applicants will be appointed, and those who are appointed will be generally appointed later. In addition, lower written scores seem to discourage applicants from pursuing the job of Firefighter, as shown below. The dramatic difference in average scores for the written test, in particular, will act to drive the rankings down for black applicants overall. To the extent that seniority is a factor in making decisions about shift or other aspects of job assignments, or "acting" assignments (as when a supervisor is on vacation), or eligibility for promotional examinations, such a difference in seniority will have career-long detrimental effects on black Firefighters.

e. Opinion on Average Test Score Differences: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against

black applicants in the average scores on the written test and PPT.

9. Adverse Impact in Rank Order Placement On the List: Exam 7029 (1999)

Focusing on the passing applicants, we can ask if there was adverse impact in the ranking of black applicants who passed the exam. This is important because no one can be appointed before they are certified, and certification is in rank order starting with the applicant ranked first on the list (who is the applicant with the highest score). Therefore, I did an adverse impact analysis of the rank order placement, comparing the placement of black and white applicants in two ways: average rank, and proportion in each grouping of 1,000 ranks. I used groupings of 1,000 ranks because the NYC FD has hired, on average, about 800 Firefighters yearly since 2002, the first year that Firefighters were hired from the 1999 exam. My adverse impact analysis of rank order placement on the list follows.¹⁶

a. Average Rank on the List: Exam 7029 (1999)

I calculated the average (i.e., the mean) rank on the list for black and white applicants based on an Excel file provided to me titled "7029 Spreadsheet FINAL.xls". The average ranks on the list are presented in Table 9a below.

Table 9a. Average Rank on the List, by Race	
Group	Mean Rank*
Black	3,813
White	3,183

* Lower numbered ranks are certified before higher numbered ranks.

b. Probability and Statistical Significance of Average Rank on the List: Exam 7029

The difference between the average rank on the list for black and for white applicants is 630 ranks, and this difference is highly statistically significant (see Table 9b). A difference this large would occur by chance less than one time in ten thousand. This corresponds to 6.5 standard deviation units.

Table 9b. Average Rank on the List, by Race: Statistical Significance					
Black Average Rank	White Average Rank	Difference	Probability	F Statistic*	Standard Deviation Units
3,813	3,183	630	< .0001	42.6	6.5

* With 1 and 5,660 d.f.

¹⁶I replicated this analysis for males only, and the results were virtually unchanged as to adverse impact and statistical significance.

c. Practical Import of Average Rank on the List: Exam 7029 (1999)

The practical importance of black applicants being ranked lower on the list is that, other things being equal, black applicants will be hired later than white applicants. In addition, lower written scores seem to discourage applicants from pursuing the job of Firefighter. The next analysis will show the lower rankings in terms of percentages within groupings of 1,000 ranks.

d. Ranks on the List in Groupings of One Thousand: Exam 7029 (1999)

I divided the list into nominal groupings of 1,000 ranks. Table 9c shows the number and percent of black and white applicants in each grouping of ranks. White applicants are uniformly spread across the groupings of ranks, with about 15% in each grouping, while black applicants tend to cluster in the bottom groupings of ranks. The top two groupings of 1,000 ranks include 30.4% of the white applicants but only 19.1% of the black applicants. The bottom two groupings include 20.7% of the white applicants but 30% of the black applicants, mirroring the totals in the top two ranks, but in reverse. In short, black applicants tend to appear lower on the list.

e. Probability and Statistical Significance of Rankings: Exam 7029 (1999)

The difference just noted in the proportion of black and white applicants between groupings of one thousand ranks is statistically significant ($\chi^2 = 42.2$, $df = 6$, $p < .0001$). A difference this large would occur by chance less than 1 time in ten thousand.

f. Practical Import of Rankings: Exam 7029 (1999)

The practical importance of black applicants being clustered in the lower rankings is that, other things being equal, black applicants will be hired later than white applicants. To the extent that seniority is a factor in making decisions about shift or other aspects of job assignments, or "acting" assignments (as when a supervisor is on vacation), or eligibility for promotional examinations, then a difference in seniority will have career-long detrimental effects on black Firefighters.

g. Opinion on Ranking on the List: Exam 7029 (1999)

Of the applicants who pass the exam, there is practically important and highly statistically significant adverse impact against black applicants in placement on the rank order eligible list.

Table 9c. Number and Percent in Each Thousand Ranks, by Race			
Grouping of Ranks	White: Number in Grouping (percent)	Black: Number in Grouping (percent)	Total of Black and White
First 1,000 ranks	796 (15.1%)	32 (8.6%)	828* (100%)
Second 1,000 ranks	810 (15.3%)	39 (10.5%)	849* (100%)
Third 1,000 ranks	878 (16.6%)	58 (15.5%)	936* (100%)
Fourth 1,000 ranks	890 (16.8%)	50 (13.4%)	940* (100%)
Fifth 1,000 ranks	819 (15.5%)	79 (21.2%)	898* (100%)
Sixth 1,000 ranks	804 (15.2%)	82 (22.0%)	886* (100%)
Seventh 1,000 ranks	292 (5.5%)	33 (8.8%)	325** (100%)
All Ranks	5,289	373	5,662

* This number is less than 1,000 because there is a total of 797 applicants in other racial or ethnic groups or who did not report an racial or ethnic group.

** There are less than 1,000 ranks in this last grouping.

10. Adverse Impact in Appointment Among Those Certified: Exam 7029 (1999)

Applicants who pass both the written test and the PPT are placed on an eligible list in rank order of their final grade. As the FD has job openings to fill, applicants from the eligible list are certified and thereby considered for appointment. The adverse impact that is seen in appointment may be due to: (a) applicant action, such as not reporting for an interview or otherwise voluntarily dropping out of the selection process or (b) some other post-examination determination made by the City, such as ineligibility for appointment due to age, or not having a drivers license or 30 college credits. To focus on the determinations made by the City, I excluded from this analysis applicants with disposition action codes that indicate they could not be appointed for reasons that appear to be voluntary, such as candidates who did not report for an interview, as in my first analysis of appointments. As opposed to my analysis summarized in Tables 1a and 1b which considered appointments from among all applicants who took the written test, this analysis considers only appointments from among applicants who passed both the written test and the PPT and were certified for consideration for appointment.

a. Number Appointed Among Those Certified: Exam 7029 (1999)

Among those applicants who were certified, I tallied the number of applicants appointed based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". The inclusions and exclusions based on disposition action codes for this analysis are the same as for my earlier analyses of adverse impact in appointment. The tallies are presented in Table 10a below, with the associated percents.

Table 10a. Appointment Among Those Certified from Exam 7029, by Racial Group			
Group	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
Black	115 (44.4%)	144 (55.6%)	259 (100%)
White	2,986 (70.1%)	1,271 (29.9%)	4,257 (100%)
Total	3,101	1,415	4,516

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

b. Adverse Impact Ratio Appointed Among Certified: Exam 7029 (1999)

As shown in Table 10b, the adverse impact ratio for appointments among those applicants who were certified is .63, which I got by dividing the percentage of blacks appointed (44.4%) by the percentage of whites appointed (70.1%), or that black

applicants were appointed at less than 2/3 the rate of white applicants. This result tells us that, of those certified, black applicants were appointed at a rate only 63% of the rate that white applicants were appointed. A statistical analysis of this adverse impact ratio is reported in Table 10b.

Table 10b. Adverse Impact In Appointment Among Those Certified from Exam 7029						
Black: Appointed*	White: Appointed	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
44.4%	70.1%	.63	< .0001	74.0	8.6	63

* All applicants offered a job are considered appointed for the sake of this analysis.

c. Probability and Statistical Significance of Appointments: Exam 7029 (1999)

The adverse impact ratio of .63 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 10b, the column titled Probability). This probability corresponds to 8.6 standard deviation units.

d. Practical Import Appointments Among Those Certified: Exam 7029 (1999)

The practical importance of hiring 44.4% of black applicants who were certified versus 70.1% of white applicants who were certified is obvious: there is disproportionate hiring, with a shortfall of 63 black appointments. This shortfall is particularly important as it comes late in the selection process. This adverse impact ratio (.63) also fails the .80 rule of thumb of the *UGESP*. Among those certified, the City hired white applicants at about 1½ times the rate of black applicants.

e. Opinion on Appointments Among Those Certified: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in appointment of males.

11. Adverse Impact in Appointment Date: Exam 7029 (1999)

Even among those appointed as Firefighters, it is possible that black applicants were appointed later than white applicants. If so, this could have weighty long term consequences for Firefighters, as well as the obvious short term consequences of being on the payroll less time. To the extent that job assignments and promotions are affected by seniority, being appointed late has serious, career-long adverse effects. Therefore, I did an analysis of the appointments by date of appointment, as follows.¹⁷

¹⁷Unlike my previous analyses of appointments, this analysis considered only people who were actually appointed, excluding people who were offered a job and declined it. This is because the data file that contains appointment date does not include such a date for people who were not appointed.

a. Date of Appointment: Exam 7029 (1999)

I tallied the applicants by date of appointment based on data in an Excel file provided to me titled "CID_DOJ_7029_qualapptd.xls". Of the 28 dates of appointment during the life of this list, 14 involved a total of 20 appointed applicants, and another 14 involved a total of 2,901 appointed applicants. I did not analyze the 20 applicants hired on 14 dates because they are such a small number. I restricted my analysis to the 2,901 applicants appointed on the other 14 dates. The dates of appointment of these 2,901 applicants are shown in Table 11a, by racial group.

Table 11a. Number of Appointments by Appointment Date: Exam 7029			
Appointment Date	Black: Number (percent)	White: Number (percent)	Total: Number (percent)
04-FEB-01	2 (1.6%)	125 (98.4%)	127 (100.0%)
06-MAY-01	1 (1.1%)	88 (98.9%)	89 (100.0%)
15-JUL-01	1 (.7%)	133 (99.3%)	134 (100.0%)
28-OCT-01	10 (3.5%)	275 (96.5%)	285 (100.0%)
27-JAN-02	12 (4.5%)	254 (95.5%)	266 (100.0%)
06-MAY-02	8 (3.0%)	263 (97.0%)	271 (100.0%)
28-JUL-02	14 (5.5%)	242 (94.5%)	256 (100.0%)
02-FEB-03	6 (2.5%)	236 (97.5%)	242 (100.0%)
04-MAY-03	10 (3.1%)	311 (96.9%)	321 (100.0%)
14-SEP-03	13 (6.2%)	196 (93.8%)	209 (100.0%)
10-DEC-03	6 (2.4%)	243 (97.6%)	249 (100.0%)
07-MAR-04	4 (1.9%)	208 (98.1%)	212 (100.0%)
25-MAY-04	8 (5.9%)	127 (94.1%)	135 (100.0%)
12-SEP-04	8 (7.6%)	97 (92.4%)	105 (100.0%)

b. Adverse Impact in Appointment Date: Exam 7029 (1999)

Black applicants tended to be appointed later than white applicants. In the first 3 appointment dates listed, only about 1% of the appointments were of black applicants (4 are black out of the 350 appointed on the first 3 appointment dates).

c. Probability and Statistical Significance of Appointment Date: Exam 7029 (1999)

The differences in proportion of black applicants appointed by appointment date, as just described, are statistically significant (Chi Sq = 25.1, d.f. = 13, $p < .02$). A difference this large would occur by chance less than 2 times in a hundred. In addition, I undertook an analysis of the mean (i.e., average) time from exam to appointment, by racial group. White applicants were appointed, on average, 81 days earlier than black applicants (see Table 11b). This difference is statistically significant. A difference this large would occur by chance less than 4 times in a hundred. This corresponds to 2.1 standard deviation units.¹⁸

Table 11b. Mean Time From Exam to Appointment: Exam 7029		
Group	Mean	Number Appointed
Black	1,440 days	104
White	1,359 days	2,817
Total	1,362 days	2,921

d. Practical Import of Appointment Date: Exam 7029 (1999)

The practical importance of black applicants being appointed later than white applicants depends on the way seniority is used in the NYC FD. To the extent that seniority is a factor in making decisions about shift or other aspects of job assignments, or “acting” assignments (as when a supervisor is on vacation), such a difference in seniority will have career-long detrimental effects on black Firefighters. The applicants who were appointed later during the life of the list waited as much as three years longer to be appointed than those appointed earlier.

e. Opinion on Appointment Date: Exam 7029 (1999)

There is practically important and statistically significant adverse impact against black applicants in appointment date.

¹⁸The statistical test employed was the F test, which is appropriate when testing means. The Chi Square statistic is not appropriate for testing the difference between means. The F value is 4.6, with 1 and 2,919 d.f.

12. Adverse Impact in CNS Disposition Action Codes: Exam 7029 (1999)

A post-certification step in the screening process involves the "Personnel Review Board" (PRB). It is my understanding that the PRB was asked to review the files of some certified applicants and, depending on the action of the PRB, some certified applicants were not hired, and these applicants are identified in the data file by the disposition action code CNS (considered not selected). As I understand it, the disposition code CNS was given to every applicant reviewed by the Personnel Review Board (PRB) and not appointed. (In the data files I have, there is no indication of which applicants were reviewed by the PRB and then appointed, nor of which applicants had their names removed from the eligible list due to the action of the PRB.) The relevant data file contains disposition action codes for the first and last certifications only. There may have been other CNS disposition action codes for other certifications, but I have no information about that.¹⁹ Based on the data available me, I did an adverse impact analysis of the PRB decisions not to select from the 1999 exam, as follows.

a. Number of CNS Disposition Action Codes: Exam 7029 (1999)

I tallied the number of certified applicants with CNS disposition action codes for either the first certification or the last certification (or both) based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS", and the total number of certified applicants. These tallies are presented in Table 12a, with the associated percents.

Table 12a. CNS Disposition Code and Racial Group: Exam 7029		
Group	Number of Applicants With CNS Disposition (percent)	Total Number of Certified Applicants (percent)
Black	18 (23.1%)	297 (6.2%)
White	60 (76.9%)	4,498 (93.8%)
Total	78 (100%)	4,795 (100%)

b. Adverse Impact in CNS Disposition Codes: Exam 7029 (1999)

A disproportionate number of black applicants were reviewed by the PRB and not hired as a result (i.e., received a CNS disposition code). Black applicants comprise 6.2% of all certified applicants so, other things being equal, we would expect 6.2% of the applicants with a CNS code to be black, but we see that 23.1% of the applicants with a

¹⁹If I receive additional data, I may redo this analysis.

CNS code are black.

c. Probability and Statistical Significance in CNS Codes: Exam 7029 (1999)

The greater proportion of black than white applicants with CNS disposition codes is highly statistically significant. A disparity this large would occur by chance less than one time in ten thousand. The Chi Square value is 36, which corresponds to 6 standard deviation units.

d. Practical Import of CNS Codes: Exam 7029 (1999)

In addition to the disproportionate number of black applicants reviewed and not selected by the PRB, of applicants who completed the screening, as previously described, applicants without a CNS disposition code were 9 times more likely to be appointed over the life of the list (69.7% versus 7.7%, see Table 12b)²⁰.

Table 12b. CNS Disposition Code by Appointment: Exam 7029			
Code	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
No CNS Code	3,095 (69.7%)	1,343 (30.3%)	4,438 (100%)
CNS Code	6 (7.7%)	72 (92.3%)	78 (100%)
Total	3,101	1,415	4,516

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

e. Opinion on CNS Disposition Codes : Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in the CNS disposition codes.

13. City's Evaluation of the Adverse Impact in Passing the Written Test: 1999 Exam

Titled "Passmarks and gender and ethnic information," Exhibit 26 shows the City's tallies of the number of applicants who would pass at various grades, with associated percents, broken down by race. I have extracted data for black and white applicants, and have computed the adverse impact ratio and conducted a statistical test for each passmark shown in Exhibit 26. This is summarized in Table 13. The first three columns of Table 13 are taken directly from Exhibit 26.

²⁰This difference is statistically significant, with a Chi Square of 134.3, which corresponds to 11.6 standard deviation units.

Table 13. Adverse Impact Based on City's Exhibit 26, Written Test 7029 (1999)						
Passmark	White: % Passing	Black: % Passing	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units
95.294	51.97	19.49	0.38	< .0001	640.7	25.3
94.117	60.37	24.48	0.41	< .0001	793.3	28.2
92.941	67.04	29.58	0.44	< .0001	909.3	30.2
91.764	72.54	34.57	0.48	< .0001	1,005.4	31.7
90.588	77.06	39.62	0.51	< .0001	1,069.6	32.7
89.411	80.79	44.55	0.55	< .0001	1,107.9	33.3
88.235	83.96	49.71	0.59	< .0001	1,108.1	33.3
87.058	86.24	54.06	0.63	< .0001	1,084.1	32.9
85.882	88.14	57.54	0.65	< .0001	1,085.2	32.9
84.705	89.84	61.19	0.68	< .0001	1,062.2	32.6

a. Adverse Impact Ratios for Passing Written Test, From Exhibit 26: Exam 7029 (1999)

The most severe adverse impact ratio for the passmarks in Exhibit 26 is .38, for the highest passmark. The adverse impact ratios are less severe with each decreasing passmark. The least severe adverse impact ratio is .68 for the lowest passmark. This clear pattern mirrors what is seen nationwide with this type of examination. These ratios tell us that the adverse impact on black applicants is greatest at the highest passmark. A statistical analysis of each of these adverse impact ratios is also reported in Table 13.

b. Probability and Statistical Significance of Adverse Impact Ratios from Exhibit 26: Exam 7029 (1999)

All 10 adverse impact ratios are highly statistically significant (see Table 13, the columns titled probability). A ratio as small as any of these would occur by chance much less than one time in ten thousand. The statistical significance for these 10 adverse impact ratios corresponds to from 25.3 to 33.3 standard deviation units.

c. Practical Import of Adverse Impact Ratios from Exhibit 26: Exam 7029 (1999)

The practical importance of these adverse impact ratios for the written test is clear: other things being equal, if relatively fewer black applicants pass the written test, then relatively fewer black applicants will be hired. At all the passmarks reported by the City in Exhibit 26, the City reports that black applicants passed the written test at about

2/5 to 2/3 the rate of white applicants, depending on the passmark. All these adverse impact ratios fail the .80 rule of thumb of the *UGESP*. In addition, it is clear that using this test to rank applicants will result in additional adverse impact among passing applicants, due to ranking.

d. Opinion on Written Test Adverse Impact Ratios from Exhibit 26: Exam 7029 (1999)

There is practically important and highly statistically significant adverse impact against black applicants in passing the written test at each of the ten passmarks reported by the City in Exhibit 26.

Section Three: Adverse Impact Analyses: 2002 Exam

14. Adverse Impact in Appointments by Race: Exam 2043 (2002)

The most basic evaluation of adverse impact involves an evaluation of the number and percent of black and white applicants appointed (or hired). Although appointments may still be made from this exam, I did an adverse impact analysis of the appointments resulting from the 2002 exam as of the effective date of the data file provided by the City, as follows.

a. Number of Appointments: Exam 2043 (2002)

I tallied the number of persons appointed based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". For the sake of this analysis, and other similar analyses below, I included as appointed those applicants with first or last disposition action codes that indicated they were: appointed, appointed from a promotional list, offered a job and turned it down, or failed to report after accepting appointment (disposition action codes APP, AOL, DEA, and FRA).²¹ For the sake of this analysis, and other similar analyses below, I excluded those applicants with disposition action codes that indicate they could not be appointed for reasons that were generally of a voluntary nature. The excluded disposition codes indicated that an applicant: was deceased, failed to report for investigation, failed to report for interview, was not qualified for appointment, was overage for the position, or was underage at time of appointment (disposition action codes DCE, FRI, FTR, NQA, OVA, UNA). These exclusions were based on a combination of the first and last disposition code: the first disposition code had to be one of those just listed, and the last disposition code had to be either one of those just listed or blank (indicating the applicant was not certified again).²²

²¹The data file available to me contained only disposition action codes for the first and last certifications. Although these are arguably the most important certifications, if and when I get disposition code information for any other certifications, I may redo this analysis.

²²I replicated this analysis counting only appointments (disposition code APP), and without the exclusions just noted, and the results were virtually unchanged as to adverse impact

These tallies are presented in Table 14a below, with the associated percents. This table shows 7.3% of the black applicants were appointed, as compared with 16.5% of the white applicants.

Table 14a. Appointments from Exam 2043, by Racial Group			
Group	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
Black	101 (7.3%)	1,277 (92.7%)	1,378 (100%)
White	2,254 (16.5%)	11,443 (83.5%)	13,697 (100%)
Total	2,355	12,720	15,075

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

b. Adverse Impact Ratio for Appointments by Race: Exam 2043 (2002)

As shown in Table 14b, the adverse impact ratio for appointments is .44, which I got by dividing the percentage of blacks appointed (7.3%) by the percentage of whites appointed (16.5%). This result tells us that black applicants were appointed at a rate only 44% of the rate that white applicants were appointed, or that black applicants had less than 1/2 the chance of being appointed as white applicants. A statistical analysis of this adverse impact ratio is reported in Table 14b.

Table 14b. Adverse Impact: Appointments from Exam 2043						
Black: Percent Appointed*	White: Percent Appointed*	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
7.3%	16.5%	.44	< .0001	78.4	8.9	114

* All applicants offered a job are considered appointed for the sake of this analysis.

c. Probability and Statistical Significance of Appointments: Exam 2043 (2002)

The adverse impact ratio of .44 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 14b, the column titled "Probability"). This probability corresponds to 8.9 standard deviation units.

and statistical significance. This analysis is summarized in Attachment E.

d. Practical Import of Appointments: Exam 2043 (2002)

The practical importance of hiring 7.3% of black applicants versus 16.5% of white applicants is obvious: there is disproportionate hiring, with a shortfall of 114 black appointments. The adverse impact ratio (.44) also fails the .80 rule of thumb of the *UGESP*. The City hired white applicants at less than 1/2 the rate of black applicants.

e. Opinion on Appointments: Exam 2043 (2002)

There is practically important and highly statistically significant adverse impact against black applicants in appointments.

15. Adverse Impact in Appointments of Males by Race: Exam 2043 (2002)

Since very few female applicants were appointed, and there were relatively more black female than white female applicants, I wanted to be sure the analysis just described was not inadvertently measuring adverse impact based on gender rather than strictly on race. Including the same disposition codes as indicating appointment, and using the same exclusions as in the analysis above, 5.2% of the female applicants were appointed while 15.8% of the male applicants were appointed; and 10.5% of the black applicants were female, while only 2.0% of the white applicants were female. Therefore I did an analysis for male applicants only. I did not do any adverse impact analyses for appointment of female applicants, because only 22 females were appointed which is a relatively small number of appointees. The adverse impact analysis of appointments from the 2002 exam considering only male applicants follows.

a. Number of Males Appointed from Exam 2043 (2002)

I tallied the number of males appointed based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls" as described in the previous analysis. These tallies are presented in Table 15a below, with the associated percents.

Table 15a. Males Appointed from Exam 2043, by Racial Group			
Group	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
Black	100 (8.1%)	1,451 (92.8%)	1,230 (100%)
White	2,233 (16.7%)	9,452 (76.0%)	13,400 (100%)
Total	2,333	10,903	14,630

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

b. Adverse Impact Ratio for Males Appointed from Exam 2043 (2002)

As shown in Table 15b, the adverse impact ratio for appointments is .49, which I got by dividing the percentage of blacks appointed (8.1%) by the percentage of whites appointed (16.7%). This result tells us that black male applicants were appointed at about 50% the rate that white male applicants were appointed, or that black male applicants were appointed at less than 1/2 the rate of white male applicants. A statistical analysis of this adverse impact ratio is reported in Table 15b.

Table 15b. Adverse Impact: Males Appointed from Exam 2043 (2002)						
Black: Percent Appointed	White: Percent Appointed	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
8.1%	16.7%	.49	< .0001	60.6	7.8	96

c. Probability and Statistical Significance of Male Appointments: Exam 2043 (2002)

The adverse impact ratio of .49 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 15b, the column titled Probability). This probability corresponds to 15.1 standard deviation units.

d. Practical Import of Male Appointments: Exam 2043 (2002)

The practical importance of hiring 8.1% of black males versus 16.7% of white males is obvious: there is disproportionate hiring, with a shortfall of 96 black male appointments. The adverse impact ratio (.49) also fails the .80 rule of thumb of the *UGESP*. The City hired white males at about 2 times the rate of black males.

e. Opinion on Male Appointments: Exam 2043 (2002)

There is practically important and highly statistically significant adverse impact against black applicants in appointment of males.

16. Adverse Impact in Passing the Written Test: Exam 2043 (2002)

There are many possible reasons for the adverse impact just shown in the appointment of black firefighters, ranging from failing the examination, to low scores on the written test and PPT that comprise the examination, to low placement on the eligible list, to differential impact of the City's post-exam processing and screening, to applicants' decisions to drop out of contention. In the next several analyses, we will go through the examination process basically in chronological order of its administration and the subsequent use of the eligible list. To pass the examination an applicant must pass both the written test (with a score of 70) and pass the PPT. I did an adverse impact analysis of the pass-fail decisions for the

written test of the 2002 exam.²³

a. Number Passing and Failing the Written Test: Exam 2043 (2002)

I tallied the number of applicants passing and failing the written test alone based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". These tallies are presented in Table 16a below, with the associated percents.

Table 16a. Pass-Fail Written Test of Exam 2043, by Racial Group			
Group	Number Pass Written (percent)	Number Fail Written (percent)	Total (percent)
Black	1,190 (85.4%)	203 (14.6%)	1,393 (100%)
White	13,495 (97.2%)	382 (2.8%)	13,877 (100%)
Total	14,685	585	15,270

b. Adverse Impact Ratio for Passing Written Test: Exam 2043 (2002)

As shown in Table 16b, the adverse impact ratio for passing the written test is .88, which I got by dividing the percentage of blacks passing the written test (85.4%) by the percentage of whites passing (97.2%). This result tells us that black applicants passed the written test at a rate of 88% of the rate that white applicants pass. Stated differently, less than 3% of the white applicants failed the written test, while 14.6% of the black applicants failed the test, more than 5 times the failure rate of white applicants. A statistical analysis of this adverse impact ratio is reported in Table 16b.

²³Sometimes the pass-fail point has little practical import, as in cases where there are so many applicants who passed that only high scoring applicants are considered for appointment. That may be the case for the 2002 exam. However, almost universally the least adverse impact is found at the passing point, so if there is adverse impact at the passing point, there certainly will be adverse impact at any higher effective passing point. I therefore did the most conservative analysis.

Table 16b. Adverse Impact: Pass-Fail Written Test of Exam 2043						
Black: Pass Written	White: Pass Written	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
85.4%	97.2%	.88	< .0001	476.9	21.8	150

c. Probability and Statistical Significance of Passing Written Test: Exam 2043 (2002)

The adverse impact ratio of .88 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 16b, the column titled Probability). This probability corresponds to 21.8 standard deviation units.

d. Practical Import of Passing Written Test: Exam 2043 (2002)

The practical importance of 85.4% of black applicants passing the written test versus 97.2% of white applicants is that there is disproportionate passing of this hurdle, with a shortfall of 150 black passers. Although the adverse impact ratio does not fail the .80 rule of thumb of the *UGESP* the shortfall is of practical significance. Only 79 black applicants²⁴ were appointed as entry level Firefighters as of the last third of 2007.²⁵ Therefore, a shortfall of 150 black passing applicants is of great practical significance.

e. Opinion: Pass-Fail Written Test: Exam 2043 (2002)

There is practically important and highly statistically significant adverse impact against black applicants in passing the written test.

17. Adverse Impact in Passing the Physical Performance Test: Exam 2043 (2002)

In addition to passing the written test, an applicant must pass the Physical Performance Test PPT (with a score of 75) to pass the exam. I did an adverse impact analysis of the pass-fail decisions²⁶ for the PPT of the 2002 exam, as follows.

²⁴Elsewhere I report 101 black applicants were appointed. But there I included as appointed 22 black applicants who were offered appointment but declined (and so were not actually appointed).

²⁵I was not told the effective date of the data file I received. I infer this time period from the date of the last large certification on the data file, 8/2/2007, and the 208 appointments were made from the 498 applicants certified on that date.

²⁶Sometimes the pass-fail point has little practical import, as in cases where there are so many applicants who pass that only high scoring applicants are considered for appointment. That was not the case for the 2002 PPT: at least 2 applicants with the lowest passing score on the PPT were appointed.

a. Number Passing and Failing the PPT: Exam 2043 (2002)

I tallied the number of applicants passing and failing the PPT alone based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". The tallies are presented in Table 17a below, with the associated percents.

Table 17a. Pass-Fail PPT of Exam 2043, by Racial Group			
Group	Number Pass PPT (percent)	Number Fail PPT (percent)	Total (percent)
Black	459 (80.1%)	114 (19.9%)	573 (100%)
White	6,067 (81.9%)	1,337 (18.1%)	7,404 (100%)
Total	6,526	1,451	7,977

b. Adverse Impact Ratio for Passing the PPT: Exam 2043 (2002)

As shown in Table 17b, the adverse impact ratio for passing the PPT is .98, which I got by dividing the percentage of blacks passing the PPT (80.1%) by the percentage of whites passing the PPT (81.9%). This result tells us that black applicants passed the PPT at a rate of 98% of the rate that white applicants did. A statistical analysis of this adverse impact ratio is reported in Table 17b.

Table 17b. Adverse Impact: Pass-Fail PPT of Exam 2043						
Black: Pass PPT	White: Pass PPT	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
80.1%	81.9%	.98	> .05	1.1	1.0	10

c. Probability and Statistical Significance of Passing the PPT: Exam 2043 (2002)

The adverse impact ratio of .98 is not statistically significant. A ratio this small would occur by chance more than 5 times in one hundred.

d. Practical Import of Passing PPT: Exam 2043 (2002)

The practical importance of 80.1% of black applicants passing the PPT versus 81.9% of white applicants is that blacks pass at the lower rate. Although the adverse impact ratio does not fail the .80 rule of thumb of the *UGESP*, the shortfall is of some practical significance. The shortfall of 10 is relatively large as compared with the total of 79

black applicants appointed from this exam.²⁷

e. Comment on Pass Rates for the PPT in 1999 and 2002 Exams

I found adverse impact in passing the PPT in the 1999 exam and not in the 2002 exam. This may be due, in part, to a higher passing rate for the PPT in 2002. (In general, there is less adverse impact seen with higher passing rates.) In 2002 the overall pass rate for the PPT for blacks and whites combined was 81.8% (6,526 passed out of 7,977, as shown in the Table 17a). In 1999, the overall pass rate for the PPT was 76.8% (5,662 passed out of 7,375, as shown in Table 4a above). This difference in passing rates is statistically significant (Chi Square = 59.4, d.f. = 1, $p < .0001$, corresponding to 7.7 standard score units).

f. Opinion on Passing the PPT: Exam 2043 (2002)

There is a practically important but statistically insignificant adverse impact against black applicants in passing the PPT.

18. Adverse Impact in Males Passing the Physical Performance Test: Exam 2043 (2002)

For reasons similar to those described for the 1999 exam, I did an adverse impact analysis of the PPT from the 2002 exam for male applicants only, as follows.

a. Number of Males Passing and Failing the PPT: Exam 2043 (2002)

I tallied the number of male applicants passing and failing the PPT alone based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". These tallies are presented in Table 18a below, with the associated percents. A statistical analysis of this adverse impact ratio is reported in Table 18b.

²⁷Elsewhere I report 101 black applicants appointed. But there I include as appointed 22 black applicants who were offered appointment but declined (and so were not actually appointed).

Table 18a. Males Passing the PPT of Exam 2043, by Racial Group			
Group	Number Pass PPT (percent)	Number Fail PPT (percent)	Total (percent)
Black	445 (83.8%)	86 (16.2%)	531 (100%)
White	6,012 (82.4%)	1,280 (17.6%)	7,292 (100%)
Total	6,457	1,366	7,823

Table 18b. Adverse Impact: Males Passing the PPT of Exam 2043						
Black: Pass PPT	White: Pass PPT	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
83.8%	82.4%	1.02	>.05	0.5	0.7	n/a

b. Results and Interpretation for Males Passing the PPT: Exam 7029 (1999)

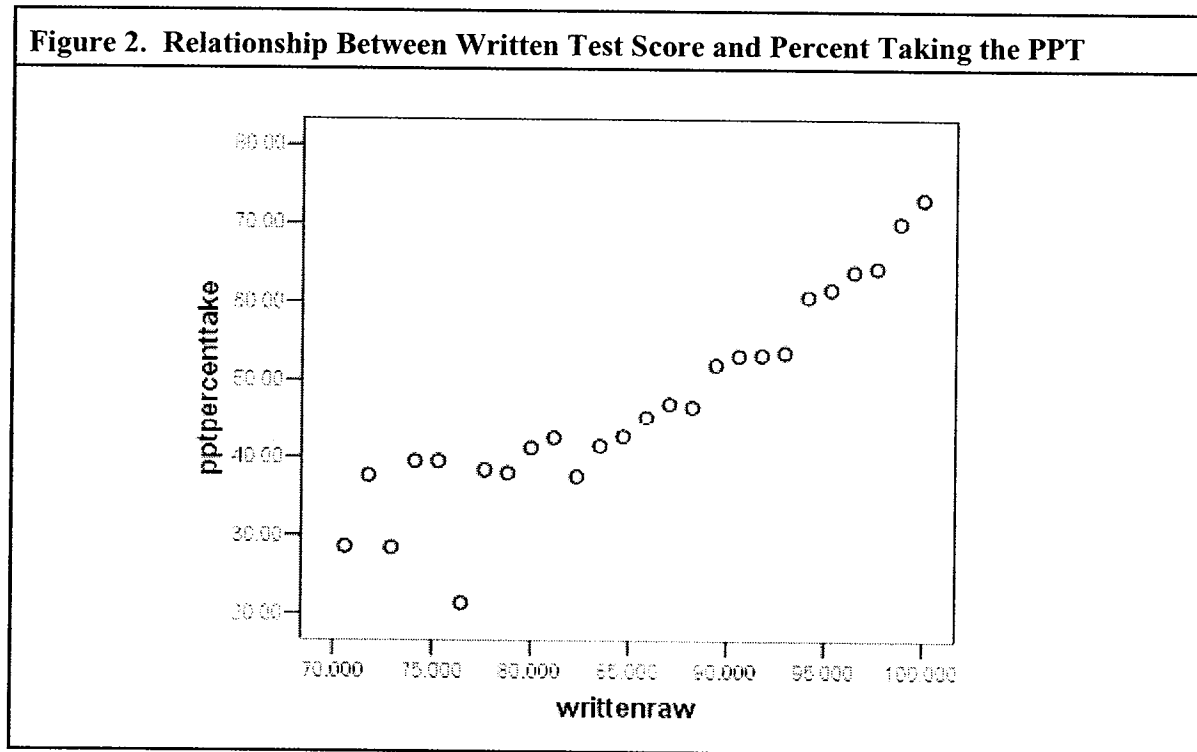
The adverse impact ratio equals 1.02, so there is no adverse impact against blacks and no shortfall.

19. Relationship Between Written Test Score and Taking the PPT: Exam 2043 (2002)

All applicants who passed the written test were allowed to take the PPT, but not all did so. A tally of the applicants who passed the written test shows that proportionally fewer black applicants took the PPT than white applicants (48.2% as compared to 54.8%, see Table 19²⁸). This reflects the general pattern that applicants who score lower on the written test are less likely to take the PPT. The clear correlation between score on the written test and the percent of applicants taking the PPT may be seen in Figure 2. The relationship between two such scores may be described in statistical terms by the "correlation coefficient," a statistic that can range from minus 1 to plus 1, with zero indicating no correlation and 1 indicating a perfect relationship. In the current case, the correlation is .933, which is extremely high, and also statistically significant ($p < .001$).

²⁸This difference is statistically significant (Chi Square = 103.0, d.f. = 4, $p < .001$).

Table 19. Applicants Taking the PPT, by Race		
Group	Number Taking PPT (percent)	Total Number Passing the Written Test
Black	573 (48.2%)	1,190
White	7,402 (54.8%)	13,495
Total	7,975 (54.3%)	14,685



Notes: Pearson correlation, $r = .933$, $p < .001$.

Writtenraw = score on written test

pptpercenttake = the percent of applicants at any written test score taking the PPT

a. Practical Import of the Correlation Between Written Test Score and Taking the PPT: Exam 2043 (2002)

The relationship described above strongly suggests that the reason proportionally fewer black applicants took the PPT than white applicants is that black applicants generally received lower scores on the written test. So the (passing) written test scores affected

the PPT scores in a way that goes beyond the intended examination weights. If black applicants had gotten higher scores on the written test, we would expect more black applicants to have taken the PPT and been hired.

b. Opinion on the Correlation Between the Written Test Score and Taking the PPT: Exam 2043 (2002)

There is a practically important and highly statistically significant relationship between the written test score and taking the PPT that serves to deter black applicants from taking the PPT.

20. Adverse Impact in Passing Overall Exam: Exam 2043 (2002)

To be included on the eligible list, which is required to be considered for appointment, an applicant must pass the overall exam (both the written test and the PPT).²⁹ My adverse impact analysis of the pass-fail decisions for the 2002 exam follows.

a. Number Passing and Failing Overall Exam 2043 (2002)

I tallied the number of applicants passing and failing the exam based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". These tallies are presented in Table 20a below, with the associated percents.

Table 20a. Pass-Fail Overall Exam 2043, by Racial Group			
Group	Number Pass Exam (percent)	Number Fail Exam (percent)	Total Number of Applicants (percent)
Black	459 (33.0%)	934 (67.0%)	1,393 (100%)
White	6,066 (43.7%)	7,812 (56.3%)	13,878 (100%)
Total	6,525	8,746	15,271

b. Adverse Impact Ratio for Passing Overall Exam 2043 (2002)

As shown in Table 20b, the adverse impact ratio for passing the exam is .76, which I got by dividing the percentage of blacks passing (33.0%) by the percentage of whites

²⁹Sometimes the pass-fail point has little practical import, as in cases where there are so many applicants who passed that only high scoring applicants are considered for appointment. That may be the case for the 2002 exam. However, as mentioned above, almost universally the least adverse impact is found at the passing point, so if there is adverse impact at the passing point, there certainly will be adverse impact at any higher effective passing point. I therefore did the most conservative analysis.

passing (43.7%). This result tells us that black applicants passed the exam at a rate of 76% of the rate that white applicants, or that black applicants passed the exam at about 3/4 the rate of white applicants. A statistical analysis of this adverse impact ratio is reported in Table 20b.

Table 20b. Adverse Impact: Pass-Fail Decisions of Overall Exam 2043						
Black: Pass Exam	White: Pass Exam	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
33.0%	43.7%	.76	< .0001	59.5	7.7	136

c. Probability and Statistical Significance of Passing Overall: Exam 2043 (2002)

The adverse impact ratio of .76 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 20b, the column titled Probability). This probability corresponds to 7.7 standard deviation units.

d. Practical Import of Passing Overall: Exam 2043 (2002)

The practical importance of passing 33.0% of black applicants versus 43.7% of white applicants is obvious: there is disproportionate passing, with a shortfall of 136 black passers. This adverse impact ratio (.76) is less than the .80 rule of thumb of the *UGESP*. The City passed white applicants at about 1 1/3 times the rate of black applicants.

e. Opinion on Passing Overall: Exam 2043 (2002)

There is practically important and highly statistically significant adverse impact against black applicants in passing the exam.

21. Average on Written Test and PPT: Exam 2043 (2002)

I describe here an overview of the average written test scores and PPT scores by race.

a. Average Test Scores: Exam 2043 (2002)

I calculated the mean scores (averages) shown in Table 21 based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls".

Table 21. Mean Scores, by Race and Gender: Exam 2043 (2002)					
Test Score*	Black Mean	White Mean	Probability	F Statistic (d.f.)	Standard Deviation Units
Written Test	81.9	90.4	< .0001	1,131.3 (1, 15268)	33.6
PPT**	87.7	90.0	< .0001	13.3 (1, 7975)	3.6
PPT: Males	89.0	90.2	> .05	3.7 (1, 7821)	1.9
PPT: Females	70.7	74.8	> .05	2.5 (1, 142)	1.5

* A higher score is better.

** These averages were computed only for applicants who took the PPT.

b. Adverse Impact in Average Test Scores: Exam 2043 (2002)

The adverse impact ratio described in the *UGESP* is not applicable here. To measure the adverse effect, I calculated the difference between the average test scores for black and white applicants. These differences are all in favor of white applicants.

c. Probability and Statistical Significance of Average Score Differences: Exam 2043 (2002)

All the mean score differences shown in Table 21 are in favor of white applicants. This difference between mean scores for black and white applicants on the written test is large and highly statistically significant (see Table 21, the column titled Probability), corresponding to 33.6 standard deviation units. The difference between mean scores on the PPT is statistically significant, but this appears to be due to the mix of males and females taking the PPT since when the analyses are done separately by gender the mean score differences between white and black applicants are not statistically significant, and since there are proportionally more black female applicants than white, and females score lower than males on the PPT, on average.

d. Practical Import of Average Score Differences: Exam 2043 (2002)

The practical importance of black applicants, on average, scoring lower than white applicants on the scores for both parts of the exam is that, other things being equal, relatively fewer black applicants will be appointed, and those who are appointed will be generally appointed later. In addition, lower written scores seem to discourage applicants from pursuing the job of Firefighter. The dramatic difference in average scores for the written test, in particular, will act to drive the rankings down for black applicants overall. To the extent that seniority is a factor in making decisions about

shift or other aspects of job assignments, or "acting" assignments (as when a supervisor is on vacation), such a difference in seniority will have career-long detrimental effects on black Firefighters.

e. Opinion on Average Test Score Differences: Exam 2043 (2002)

There is practically important and highly statistically significant adverse impact against black applicants in the average scores on the written test.

22. Adverse Impact in Rank Order Placement On the List: Exam 2043 (2002)

Focusing on the passing applicants, we can ask if there was adverse impact in the ranking of black applicants who passed the exam. This is important because no one can be appointed before they are certified, and certification is in rank order starting with the applicant ranked first on the list (who is the applicant with the highest score). Therefore, I did an adverse impact analysis of the rank order placement, comparing the placement of black and white applicants in two ways: average rank, and proportion in each grouping of 1,000 ranks, as in my analysis of the 1999 exam. My adverse impact analysis of rank order placement on the list follows.³⁰

a. Average Rank on the List: Exam 2043 (2002)

I calculated the average (i.e., the mean) rank on the list for black and white applicants based on an Excel file provided to me titled "7029 Spreadsheet FINAL.xls". The average ranks on the list are presented in Table 22a below.

Table 22a. Average Rank on the List, by Race	
Group	Mean Rank*
Black	4,598
White	3,624

* Lower numbered ranks are certified before higher numbered ranks.

b. Probability and Statistical Significance of Average Rank on the List: Exam 2043

The difference between the average rank on the list for black and for white applicants is 974 ranks, and this difference is highly statistically significant (see Table 22b). A difference this large would occur by chance less than one time in ten thousand. This corresponds to 9.6 standard deviation units.

³⁰I replicated this analysis for males only, and the results were virtually unchanged as to adverse impact and statistical significance.

Table 22b. Average Rank on the List, by Race: Statistical Significance					
Black	White	Difference	Probability	F Statistic*	Standard Deviation Units
4,598	3,624	974	< .0001	91.5	9.6

* With 1 and 6,523 d.f.

c. Practical Import of Average Rank on the List: Exam 2043 (2002)

The practical importance of black applicants being ranked lower on the list is that, other things being equal, black applicants will be hired later than white applicants. In addition, lower written scores seem to discourage applicants from pursuing the job of Firefighter. The next analysis will show the lower rankings in terms of percentages within groupings of 1,000 ranks.

d. Ranks on the List in Groupings of One Thousand: Exam 2043 (2002)

I divided the list into nominal groupings of 1,000 ranks. Table 22c shows the number and percent of black and white applicants in each grouping of ranks. White applicants are uniformly spread across the groupings of ranks, with some 13 to 14% in each grouping, while black applicants tend to cluster in the bottom groupings of ranks. The top two groupings of 1,000 ranks include 27.3% of the white applicants but only 18.1% of the black applicants. The bottom two groupings include 17.0% of the white applicants but 36.4% of the black applicants, mirroring the totals in the top two ranks, but in reverse. In short, black applicants tend to appear lower on the list.

e. Probability and Statistical Significance of Rankings: Exam 2043 (2002)

The difference just noted in the proportion of black and white applicants between groupings of one thousand ranks is statistically significant (Chi Sq = 138.1, df 7, $p < .0001$). A difference this large would occur by chance less than 1 time in ten thousand.

f. Practical Import of Rankings: Exam 2043 (2002)

The practical importance of black applicants being clustered in the lower rankings is that, other things being equal, black applicants will be hired later than white applicants. To the extent that seniority is a factor in making decisions about shift or other aspects of job assignments, or "acting" assignments (as when a supervisor is on vacation), then a difference in seniority will have career-long detrimental effects on black Firefighters.

g. Opinion on Ranking on the List: Exam 2043 (2002)

Of the applicants who pass the exam, there is practically important and highly statistically significant adverse impact against black applicants in placement on the rank order eligible list.

Table 22c. Number and Percent in Each Thousand Ranks, by Race			
Grouping of Ranks	White: Number in Grouping (percent)	Black: Number in Grouping (percent)	Total of Black and White
First 1,000 ranks	837 (13.8%)	32 (7.0%)	869* (100%)
Second 1,000 ranks	820 (13.5%)	51 (11.1%)	871* (100%)
Third 1,000 ranks	864 (14.2%)	42 (9.2%)	906* (100%)
Fourth 1,000 ranks	847 (14.0%)	52 (11.3%)	899* (100%)
Fifth 1,000 ranks	851 (14.0%)	46 (10.0%)	897* (100%)
Sixth 1,000 ranks	816 (13.5%)	69 (15.0%)	885* (100%)
Seventh 1,000 ranks	752 (12.4%)	102 (22.2%)	854* (100%)
Eighth 1,000 ranks**	279 (4.6%)	65 (14.2%)	344* (100%)
All Ranks	6,066	459	6,525

* This number is less than 1,000 because there is a total of 960 applicants in other racial or ethnic groups or who did not report an racial or ethnic group.

** There are less than 1,000 ranks in this last grouping.

23. Adverse Impact in Appointment Among Those Certified: Exam 2043 (2002)

Applicants who pass both the written test and the PPT are placed on an eligible list in rank order of their final grade. As the FD has job openings to fill, applicants from the eligible list are certified and thereby considered for appointment. The adverse impact that is seen in appointment may be due to: (a) applicant action, such as not reporting for an interview or otherwise voluntarily dropping out of the selection process or (b) some other post-examination determination made by the City, such as ineligibility for appointment due to age, or not having a drivers license or 30 college credits. To focus on the determinations made by the City, I excluded from this analysis applicants with disposition action codes that indicate they could not be appointed for reasons that appear to be voluntary, such as candidates who did not report for an interview, as in my first analysis of appointments. As opposed to my first analysis for Exam 2043 which considered appointments from among all applicants who took the written test, this analysis considers only appointments from among applicants who passed both the written test and the PPT and were certified for consideration for appointment.

a. Number Appointed Among Those Certified: Exam 2043 (2002)

Among those applicants who were certified, I tallied the number of applicants appointed based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". The inclusions and exclusions based on disposition action codes for this analysis are the same as for my earlier analyses of adverse impact in appointment. The tallies are presented in Table 23a below, with the associated percents.

Table 23a. Appointment Among Those Certified from Exam 2043, by Racial Group			
Group	Number Appointed* (percent)	Number Not Appointed (percent)	Total** (percent)
Black	101 (73.2%)	37 (26.8%)	138 (100%)
White	2,254 (81.9%)	499 (18.1%)	2,753 (100%)
Total	1,901	1,185	2,891

* All applicants offered a job are considered appointed for the sake of this analysis.

** Applicants who did not complete the post-exam screening are omitted from this analysis.

b. Adverse Impact Ratio Appointed Among Certified: Exam 2043 (2002)

As shown in Table 23b, the adverse impact ratio for appointments among those applicants who were certified is .89, which I got by dividing the percentage of blacks appointed (73.2%) by the percentage of whites appointed (81.8%). A statistical analysis of this adverse impact ratio is reported in Table 23b.

Table 23b. Adverse Impact In Appointment Among Those Certified from Exam 2043						
Black: Pass PPT	White: Pass PPT	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units	Shortfall
73.2%	81.9%	.89	.014	6.3	2.4	11

* All applicants offered a job are considered appointed for the sake of this analysis.

c. Probability and Statistical Significance of Appointments: Exam 2043 (2002)

The adverse impact ratio of .89 is statistically significant. A ratio this small would occur by chance much less than two times out of one hundred (see Table 23b, the column titled Probability). This probability corresponds to 2.4 standard deviation units.

d. Practical Import Appointments Among Those Certified: Exam 2043 (2002)

The practical importance of hiring 73.2% of black applicants who were certified versus 81.9% of white applicants who were certified is that there is disproportionate hiring, with a shortfall of 11 black appointments. Although the adverse impact ratio does not fail the .80 rule of thumb of the *UGESP* the shortfall is of practical significance. Only 79 black applicants³¹ were appointed as entry level Firefighters as of mid to late 2007. Therefore, a shortfall of 11 black applicants is of practical significance, particularly this late in the selection process.

e. Opinion on Appointments Among Those Certified: Exam 2043 (2002)

There is practically important and statistically significant adverse impact against black applicants in appointment among those certified.

³¹Elsewhere I report 101 black applicants were appointed. But there I included as appointed 22 black applicants who were offered appointment but declined (and so were not actually appointed).

24. Adverse Impact in Appointment Date: Exam 2043 (2002)

Even among those appointed as Firefighters, it is possible that black applicants were appointed later than white applicants. If so, this could have weighty long term consequences for Firefighters, as well as the obvious short term consequences of being on the payroll less time. To the extent that job assignments and promotions are affected by seniority, being appointed late has serious, career-long adverse effects. Therefore, I did an analysis of the appointments by date of appointment, as follows.³²

a. Date of Appointment: Exam 2043 (2002)

I tallied the applicants by date of appointment based on data in an Excel file provided to me titled "CID_DOJ_7029_qualapptd.xls". Of the 15 dates of appointment during the life of this list, 3 involved 1 appointment each, and another 12 involved a total of 1,908 appointed applicants. I did not analyze the 3 applicants hired on 3 dates because they are such a small number. I restricted my analysis to the 1,908 applicants appointed on the other 12 dates. The dates of appointment of these 1,908 applicants are shown in Table 24a, by racial group.

³²Unlike my previous analyses of appointments, this analysis considered only people who were actually appointed, excluding people who were offered a job and declined it. This is because the data file that contains appointment date does not include such a date for people who were not appointed.

Table 24a. Number of Appointments by Appointment Date: Exam 2043			
Appointment Date	Black: Number (percent)	White: Number (percent)	Total: Number (percent)
25-MAY-04	0 (.0%)	20 (100.0%)	20 (100.0%)
12-SEP-04	0 (.0%)	44 (100.0%)	44 (100.0%)
05-DEC-04	2 (1.5%)	130 (98.5%)	132 (100.0%)
08-MAR-05	12 (5.6%)	204 (94.4%)	216 (100.0%)
31-MAY-05	3 (1.4%)	210 (98.6%)	213 (100.0%)
25-SEP-05	12 (7.6%)	145 (92.4%)	157 (100.0%)
15-JAN-06	3 (2.9%)	100 (97.1%)	103 (100.0%)
11-APR-06	8 (5.2%)	147 (94.8%)	155 (100.0%)
11-JUN-06	5 (3.0%)	159 (97.0%)	164 (100.0%)
19-NOV-06	7 (3.9%)	173 (96.1%)	180 (100.0%)
25-MAR-07	17 (6.7%)	238 (93.3%)	255 (100.0%)
05-AUG-07	11 (4.0%)	261 (96.0%)	272 (100.0%)
Total	80 (4.2%)	1,831 (95.8%)	1,911 (100.0%)

b. Adverse Impact in Appointment Date: Exam 2043 (2002)

Black applicants tended to be appointed later than white applicants. In the first 3 appointment dates listed, only about 1% of the appointments were of black applicants (2 are black out of the 196 appointed on the first 3 appointment dates).

c. Probability and Statistical Significance of Appointment Date: Exam 2043 (2002)

The differences in proportion of black applicants appointed by appointment date, as just described, are statistically significant (Chi Sq = 20.2, d.f. = 11, $p < .05$). A difference this large would occur by chance less than 5 times in a hundred. In addition, I undertook an analysis of the mean (i.e., average) time from exam to appointment, by racial group. White applicants were appointed, on average, 68 days earlier than black applicants (see Table 24b). This difference is statistically significant. A difference this large would occur by chance less than 5 times in a hundred. This corresponds to 1.7 standard deviation units.³³

Table 24b. Mean Time From Exam to Appointment: Exam 2043		
Group	Mean	Number Appointed
Black	1,259 days	80
White	1,191 days	1,834
Total	1,194 days	1,914

d. Practical Import of Appointment Date: Exam 2043 (2002)

The practical importance of black applicants being appointed later than white applicants depends on the way seniority is used in the NYC FD. To the extent that seniority is a factor in making decisions about shift or other aspects of job assignments, or "acting" assignments (as when a supervisor is on vacation), such a difference in seniority will have career-long detrimental effects on black Firefighters. The applicants who were appointed later during the life of the list waited as much as three years longer to be appointed than those appointed earlier, and some applicants are still waiting.

e. Opinion on Appointment Date: Exam 2043 (2002)

There is practically important and statistically significant adverse impact against black applicants in appointment date.

25. An Evaluation of the Effect of Typos and Similar Data Errors

The data contains a few apparent typos or similar inconsistencies, as mentioned in above when I discussed the integrity of the data. In the next analysis I simulate the effect of errors in the data. To evaluate the effect of a small number of changes in the data, I created a new data set by changing the data for 50 applicants in Attachment E. If I randomly changed the data, the tallies would be expected not to change at all, or by very little. So I made all the data changes in a direction that favors the defendants. I changed 25 black applicants from

³³The statistical test employed was the t-test, which is appropriate when testing the difference between two means. The Chi Square statistic is not appropriate for testing the difference between means. The t-test allows for one-tailed tests. The t value is 1.7, with 1,912 d.f.

unappointed to appointed, and I changed 25 white applicants from appointed to unappointed. These changes go beyond the number of the apparent typos or similar inconsistencies I noted and, unlike random errors, these arbitrary changes are all in favor of the defendants. The data with these arbitrary changes are shown in Table 25, together with the original tallies from Attachment E. The 25 changes in each cell are indicated by +25 or -25, and the new (arbitrarily modified) tallies are shown after the equals sign.

Table 25a. Appointments from Exam 2043, by Racial Group (Modified)			
Group	Number Appointed (percent)	Number Not Appointed (percent)	Total (percent)
Black	79+25=104 (7.5%)	1,314-25=1,289 (92.5%)	1,393 (100%)
White	1,822-25=1,797 (12.7%)	12,056+25=12,081 (87.3%)	13,878 (100%)
Total	1,901	13,370	15,271

a. Adverse Impact Ratio

The adverse impact for appointments in Table 14a was .44, and here is it .59 (see Table 25b). Both of these fail the 80% rule.

Table 25b. Adverse Impact: Appointments from Exam 2043 (Modified)					
Black: % Appointed	White: % Appointed	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units
7.5%	12.7%	.59	< .0001	34.4	5.9

b. Statistical Significance

The adverse impact ratio seen for these arbitrarily modified data is .59, which is highly statistically significant (see Table 25b). A ratio this small would occur by chance much less than one time in ten thousand. This corresponds to 5.9 standard deviation units. For the actual data, the Chi Square statistic was 63.9, corresponding to 8.0 standard deviation units (see Attachment E).

c. Interpretation

This exercise in analyzing an arbitrarily modified set of data from one exam leads me to conclude that small numbers of typos in the data sets should not change most, if any of the adverse impact statistical analyses of the 1999 and 2002 exams.

Section Four: Census Comparisons

26. Adverse Impact of One Year of College by Race: Census 2000

I looked at data from the official 2000 Census to evaluate the adverse impact on blacks of one part of the entrance requirements listed on the examination notice: completion of 30 college credits.³⁴

a. Census Information

The official website of the US Census provided relevant comparison information on males residents of the five Boroughs of New York City, or Nassau, Orange, Putnam, Rockland, Suffolk, or Westchester Counties who are not disabled and who are between from ages 18 to 34 years of age.³⁵ There is no one page of the Census website with census information for just this grouping of people but compilation of information from various pages of the Census site yielded information that allowed an estimate of the number of black and of white people overall and who have completed one year of college.³⁶ The percents I calculated from the Census data are presented in Table 26a.

Table 26a. Education by Racial Group: 2000 Census		
Group	Less Than One Year of College	One Year of College or More
Black	59.9%	40.1%
White	34.4%	65.6%

b. Adverse Impact Ratio for One Year College: 2000 Census

As shown in Table 26b, the adverse impact ratio for requiring one year of college is .61, which I got by dividing the percentage of blacks with this level of college (40.1%) by the percentage of whites with this level of college (65.6%). A statistical analysis of this adverse impact ratio is reported in Table 26b.

³⁴This is not an absolute requirement, as individuals with a high school diploma, or equivalent, and at least two years of honorable service in the military also qualify.

³⁵This age range goes beyond the age of 29 which is the cut off stated on the examination notice. However, this age range covers almost the complete range on the examination announcement, and the census information is not available for people who precisely meet the City's entrance requirements for Firefighter.

³⁶The Census does not report college credits earned. However, one year of college is typically 30 credits.

Table 26b. Adverse Impact: Education by Racial Group, 2000 Census					
Black: Percent 1+ Years College	White: Percent 1+ Years College	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units
40.1%	65.6%	.61	< .0001	39,093.4	197.7

c. Probability and Statistical Significance for One Year College: 2000 Census

The adverse impact ratio of .61 is highly statistically significant. A ratio this small would occur by chance much less than one time in ten thousand (see Table 26b, the column titled Probability). This probability corresponds to 197.7 standard deviation units.

d. Practical Import of One Year College: 2000 Census

The practical importance of requiring one year of college is obvious: proportionally more white male residents will qualify based on their present educational achievement, without having to enroll in additional college courses. The adverse impact ratio (.61) also fails the .80 rule of thumb of the *UGESP*.

e. Opinion on One Year College: 2000 Census

There is practically important and highly statistically significant adverse impact against black male residents in requiring one year of college education.

27. Adverse Impact in Male Applications: 1999 Exam

It is possible that potential black male applicants did not apply for the job of Firefighter at the same rate as potential white male applicants.³⁷ I did an adverse impact analysis comparing the number of black and white male applicants to similar information from the 2000 Census.

a. Number of Male Applicants and Census Information

I tallied the number of males who applied based on data in an Excel file provided to me titled "exam7029 corrected applicant data.xls - (# Legal 1820012).XLS". These tallies are presented in Table 27a below, with the associated percents. The official website of the US Census provided relevant comparison information on males who met requirements on the City's Examination Notice for Firefighter: residents of the five Boroughs or Nassau, Orange, Putnam, Rockland, Suffolk, or Westchester Counties, who have 30 semester college credits, and are the age range 17.5 to 29 years of age, or are high school graduates (or GED) and served honorably for two years in the US military. There is no one page of the Census website with census information for just

³⁷I looked at males because they comprise the vast majority of the applicants and the census data for males and females is somewhat different, so a single sex comparison may be clearer.

this grouping of people but, considering only black and white male residents, compilation of information from various pages of the Census site yields ways to estimate the proportion of residents of these geographic areas that otherwise meet the requirements on the Examination Announcement. Based on this review of the Census information black males comprise 18.3% of the qualified black and white male residents in the geographic area described above.

Table 27a. Male Applicants for Exam 7029, by Racial Group		
Group	Number Applying (percent)	Percent of Qualified Residents
Black	1,601 (11.2%)	18.3%
White	12,673 (88.8%)	81.7%
Total	14,274	100.0%

b. Adverse Impact Analysis

To measure the adverse effect, if any, first I calculated the number of black and white male applicants that would be expected if the applicant group reflected the Census information. This is simply the total number of black and white male applicants (i.e., 14,274) multiplied by 18.3% or 81.7% for black and white groups, respectively. (This approach holds the total number of applicants constant. This is a typical way to statistically analyze this type of data.) Then I found the difference between the observed and the expected number of black and white male applicants. (Since the total is constant, any increase in the number of black male applicants has to be accompanied by a equal size decrease in white male applicants.) This shortfall is 1,011 black applicants, as shown in Table 27b.

Table 27b. Adverse Impact: Male Applicants for Exam 7029						
Group	Number Applying (percent)	Number Expected Based on 2000 Census (percent)	Shortfall	Probability	Chi Square*	Standard Deviation Units
Black	1,601 (11.2%)	2,612 (18.3%)	1,011	<.0001	479.1	21.9
White	12,673 (88.8%)	11,662 (81.7%)	-1,011			
Total	14,274	14,274				

* This is a "goodness of fit" Chi Square test with one degree of freedom. It may be interpreted similarly to the other Chi Square tests in this report.

c. Statistical Significance

The difference between the observed and expected numbers of black and white male applicants is highly statistically significant (see Table 27b). Differences this large would occur by chance much less than one time in ten thousand. This corresponds to 21.9 standard deviation units.

d. Practical Import

If qualified black and white male residents applied at the same rate, and if the total number of male applicants remained the same, the statistical analysis shows we would expect an additional 1,011 black applicants.

e. Opinion: Applications

There is practically important and highly statistically significant adverse impact against black male residents. It may be that black residents do not see the NYC FD as a viable employer for them.

28. Adverse Impact in Male Applications: 2002 Exam

It is possible that potential black male applicants did not apply for the job of Firefighter at the same rate as potential white male applicants. I did an adverse impact analysis comparing the number of black and white male applicants to similar information from the 2000 Census.

a. Number of Male Applicants and Census Information

I tallied the number of males who applied based on data in an Excel file provided to me titled "2043 Spreadsheet FINAL.xls". These tallies are presented in Table 28a below, with the associated percents. I include in the table the census information described in the analogous analysis for the 1999 exam.

Table 28a. Male Applicants for Exam 2043, by Racial Group		
Group	Number Applying (percent)	Percent of Qualified Residents
Black	1,711 (11.9%)	18.3%
White	12,674 (88.1%)	81.7%
Total	14,385	--

b. Adverse Impact Analysis

To measure the adverse effect, first I calculated the number of black and white male applicants that would be expected if the applicant group reflected the Census information. This is simply the total number of black and white male applicants (i.e., 14,274) multiplied by 18.3% or 81.7% for black and white groups, respectively. (This approach holds the total number of applicants constant. This is a typical way to statistically analyze this type of data.) Then I found the difference between the observed and the expected number of black and white male applicants. (Since the total is constant, any increase in the number of black male applicants has to be accompanied by a equal size decrease in white male applicants.) This shortfall is 921 black applicants, as shown in Table 28b.

Table 28b. Adverse Impact: Male Applicants for Exam 7029						
Group	Number Applying (percent)	Number Expected Based on 2000 Census (percent)	Shortfall	Probability	Chi Square*	Standard Deviation Units
Black	1,711 (11.9%)	2,632 (18.3%)	921	<.0001	394.9	19.9
White	12,674 (88.1%)	11,752 (81.7%)	-921			
Total	14,385	14,385				

* This is a "goodness of fit" Chi Square test with one degree of freedom. It may be interpreted similarly to the other Chi Square tests in this report.

c. Statistical Significance

The difference between the observed and expected numbers of black and white male applicants is highly statistically significant (see Table 28b). Differences this large would occur by chance much less than one time in ten thousand. This corresponds to 19.9

standard deviation units.

d. Practical Import

If qualified black and white male residents applied at the same rate, and if the total number of male applicants remained the same, the statistical analysis shows we would expect an additional 921 black applicants.

e. Opinion: Applications

There is practically important and highly statistically significant adverse impact against black male residents.

Section Five: Opinions Summary

29. Opinion Summary

In summary, concerning both Exam 7029 and Exam 2043 (the 1999 and 2002 entry-level Firefighter Exams), it is my professional opinion that the data provided by the City show both examinations had severe adverse impact in the appointment of black applicants, and that this adverse impact is practically and highly statistically significant.³⁸ The written test, the first hurdle in the examination, also had severe, practically important and highly statistically significant adverse impact for the 1999 exam, and less severe but still practically important and highly statistically significant adverse impact for the 2002 exam. There was practically important and statistically significant adverse impact in passing the PPT for the 1999 exam, and in taking the PPT for both the 1999 and 2002 exams. The lower written test scores for blacks, on average, led blacks to be less likely to appear to take the PPT than if they had received higher written test scores. There was practically important and highly statistically significant adverse impact in passing both the 1999 and the 2002 examinations. Further, even the black applicants appointed generally suffer from lower seniority than white applicants, something that is especially important in the fire service, and this too is statistically significant.

The City's own evaluation of the passing rates for blacks and whites of the 1999 exam, apparently completed before setting the passing score at 84.701, shows that the city was aware of the differential passing rates for blacks and whites at this passing score.

There were fewer black applicants for both the 1999 and the 2002 exams than would be expected based on the 2000 Census information on black residents of the City and Nassau, Orange, Putnam, Rockland, Suffolk, or Westchester Counties, and this is highly statistically significant for each exam. An analysis of the Census data also showed that proportionally

³⁸Some of the tests of statistical significance show more extreme values of Chi Square and probability, and related standard deviation units, than any other of the many similar tests I recall conducting over the course of my career.

fewer black residents had completed one year of college, and this difference is highly statistically significant.

Signed: _____

Date: _____

Attachment A: Documents Reviewed from Vulcan Society et al. v City of NY et al.*

My Document Number	Title of Document
1	2043 Spreadsheet FINAL
2	Exam 7029 corrected applicant data
3	Exhibit 40 Code descriptions for disposition action field of 2 exam xls files
4	Letter to NY Law Dept from Seeley, DOJ 10/12/2007
5	Exhibit 11 Exam 7029 AM session February 27, 1999
6	Exhibit 12 Exam 7029 PM session February 27, 1999
7	Exhibit 13 Exam 7029 Sabbath observer session February 26, 1999
8	Exhibit 14 Exam 7029 alternate exam (given several weeks later) March 26, 1999
9	Exhibit 15 Exam 7029 special military exam November 19, 1999
10	Exhibit 16 Exam 7029 final answer key for AM, PM, and Sabbath Observer Sessions February 26-27, 1999
11	Exhibit 17 Memory Booklet for Exam 2043
12	Exhibit 18 Exam 2043 test questions, AM session 12/14/02

My Document Number	Title of Document
13	Exhibit 19 Exam 2043 test questions, PM session 12/14/02
14	Exhibit 20 Exam 2043 test questions, Sabbath observer session 12/15/02
15	Exhibit 21 Exam 2043 test questions, special military session #1 5/16/03
16	Exhibit 22 Exam 2043 test questions, special military session #2 10/31/03
17	Exhibit 23 Exam 2043 test questions, special military session #3 3/6/05
18	Exhibit 24 Exam 2043 final answer key for all sessions
19	Exhibit 25 Summary of written test scores for Exam 7029 6/25/99
20	Exhibit 26 Passmark distribution for Exam 7029
21	Exhibit 2 Operative notice of examination for Exam 7029
22	Exhibit 4 Operative notice of examination for Exam 2043
23	Patitucci Exhibit 9 & 10 Memory booklet for exam 7029
24	Plaintiffs-Intervenors' Complaint (dated 9/25/07)
25	CID disks (sic) from the City for 2043
26	CID disks (sic) from the City for 7029

* In addition to the documents listed above, I accessed some material on the official NYC

website.

Attachment B: Resume of Joel P. Wiesen, Ph.D.

27 Judith Road	(617) 244-8859 (work)
Newton, Massachusetts 02459-1715	(617) 332-6984 (home)
Education	Ph.D. in Psychology, Lehigh University, 1975 M.A. in Psychology, C. W. Post College, 1969 B.A. in Psychology, Stony Brook State University, 1967
Employment History	Consultant in Industrial Psychology, 1977 - present Consult to public and private organizations, primarily in the areas of personnel assessment, selection, and layoff; develop tests of knowledge, ability, aptitude, and work-style; conduct applied research concerning employment in organizational settings. (Full-time since 1/94.) Director of Test Development and Validation, 1977 - 1993 Massachusetts Department of Personnel Administration Develop, implement and evaluate employee selection, promotion, and performance evaluation tools and programs; conduct applied research concerning civil service employment. (Part-time position 1987 to 1993.)
Professional Accomplishments	Plan, develop and implement programs in support of excellence and fairness in human resource management: develop culturally fair personnel assessment, selection and promotion tests and systems; conduct validation research and statistical analyses; develop diagnostic math tests for use in school settings; conduct program evaluation and redesign; expert witness in employment discrimination and sobriety testing cases; develop employee performance evaluation programs; and publish/deliver reports, papers.
Teaching Experience	Northeastern University, School of Engineering, Engineering Management Program Lecturer, 1978 - 1994 Lesley College, School of Management Adjunct Faculty, 1989 - 1994 Kings College, Psychology Department Asst. Prof. and Acting Department Chair, 1969 - 1975
Professional Associations	International Public Management Association for Human Resources Assessment Council, American Psychological Association, Association for Psychological Science, New England Society for Applied Psychology, Society of Industrial and Organizational Psychology, American Statistical Association.
Professional Honors	Past-President of IPMA-Assessment Council, Past-President of the New England Society of Applied Psychologists.
Licensure	Licensed as psychologist in two states (Massachusetts and Pennsylvania).

Attachment C: List of Papers and Publications

(1997) Adverse Impact and Testing, *NESAP News*, vol. 2, Issue 1.

(1997, June) *A New Approach to Measuring Mechanical Aptitude*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Newport Beach, CA.

(1997, June) *A Crash Course in Personnel Selection*. Pre-conference training program presented at the annual meeting of the International Personnel Management Association's Assessment Council, Newport Beach, CA. (with T. R. Lin, Ph.D. and K. Shultz, Ph.D.)

(1997, June) *Firefighter Selection With Low Adverse Impact: A Replication*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Newport Beach, CA.

(1998, May) *Development of a New Test of Mechanical Aptitude*. Invited presentation to the New England Society of Applied Psychologists (NESAP), Newton, MA.

(1999, June) *A Review and Critique of Published Aptitude Tests for Computer Programmer*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Trade Winds Resort, FL.

(1999, June) *The WTMA: a Measure of Mechanical Aptitude With Reduced Adverse Impact*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Trade Winds Resort, FL.

(2000, June) *Content-Oriented Cultural Bias Review of Test Questions*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Arlington, VA.

(2001) *Technical Manual for the Diagnostic Test of Pre-Algebra Math (DT-PAM)*. Applied Personnel Research, Newton, MA.

(2001, April) *Reducing Adverse Impact While Maintaining Validity; The Public Sector Experience: Some Possible Reasons for Adverse Impact*. Paper presented at the Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

(2001, May) *Using Tests to Minimize Adverse Impact While Maintaining Validity*. Invited presentation to the Recruitment and Hiring Seminar of the Massachusetts Chiefs of Police Association Westborough, MA.

(2001, June) *Development of a New Test for Computer Programmer/Analyst Aptitude*.

Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Newport Beach, CA.

(2001, June) *Tutorial: Review of Multiple-Choice Test Items, With Emphasis on Cultural Bias*. Presented at the annual meeting of the International Personnel Management Association's Assessment Council, Newport Beach, CA.

(2002, April) *Review of Written, Multiple-Choice Test Items, With Some Emphasis on Cultural Bias*. Paper presented at the Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Canada.

(2002, July) *Developing Defensible Written Test Questions: Art, Science, and Some Guidelines* (with I. Gast, Ph.D. and D. Hamill). Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, New Orleans, LA.

(2002, July) *The PCBS: A Quantified, Public Safety Candidate Background Self-Report Instrument*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, New Orleans, LA.

(2002, October) *Some Possible Reasons for the Black-White Mean Score Differences Seen With Many Cognitive Ability Tests*. Paper presented at the annual meeting of the Institute for the Study and Promotion of Race and Cultures, Boston, MA.

(2003) *How to Prepare for the Mechanical Aptitude and Spatial Relations Tests*. Barrons Educational Series, Hauppauge, NY.

(2003) *Technical Manual For The Diagnostic Test for High School Math (DT-HSM)*, Applied Personnel Research, Newton, MA.

(2003) *Technical Manual For The Language-Free Computer Programmer/Analyst Aptitude Test (LPAT)*, Applied Personnel Research, Newton, MA.

(2003) *Technical Manual For The Police Candidate Background Self-Report (PCBS)*, Applied Personnel Research, Newton, MA.

(2003) *Guidelines and Suggestions for Avoiding Cultural Bias in Multiple-Choice Test Questions*. Reprinted by the ERIC Clearinghouse on Counseling & Student Services (ERIC/CASS), in the 2003 ERIC/CASS publication, *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*, edited by Janet Wall and Garry Walz.

(2003, June) *An In Depth, Content Review of Spatial Ability Tests*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Baltimore, MD.

(2003, June) *The Potential of Diagnostic Math Tests*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Baltimore, MD.

(2004, October) *Adverse Impact: Theory & Practical Approaches*. Invited Address to the MAPAC Training Workshop, New York, NY.

(2005, June) Innovative, Psychometrically Sound Approaches to Hiring Firefighters, Member Showcase Presentation to the New England Society for Applied Psychology, Weston, MA.

(2005, June) *Statistical Support of Test Fairness Reconsidered*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Orlando, FL.

(2005, June) *Minimizing Adverse Impact: Learning from some Serious Flaws in the Wonderlic Personnel Test*. Paper presented at the annual meeting of the International Personnel Management Association's Assessment Council, Orlando, FL.

(June, 2006) *Thinking Outside the Box in Merit Selection*. Paper presented at the 30th Annual IPMAAC Conference on Personnel Selection, Las Vegas, NV.

(June, 2006) Benefits, Drawbacks, and Pitfalls of z-Score Weighting. Paper presented at the 30th Annual IPMAAC Conference on Personnel Selection, Las Vegas, NV.

(October, 2006) *Limitations and Flaws in the Research Supporting Field Sobriety Tests*. Invited address presented at the National Association of Criminal Defense Lawyers/National College for DUI Defense Seminar, Las Vegas, NV.

(November, 2006) *Thinking Outside the Box in Personnel Selection*. Invited address presented at the MAPAC Fall Conference, Princeton, NJ.

(June, 2007) *Possible New Approaches to Reduce Adverse Impact*. Paper presented at the 31st Annual IPMAAC Conference on Personnel Assessment, St. Louis, MO.

(June, 2007) *Tips on Writing an Expert Witness Report*. Paper presented at the 31st Annual IPMAAC Conference on Personnel Assessment, St. Louis, MO.

(June, 2007) *Introduction to Using SPSS Command Files*. Paper presented at the 31st Annual IPMAAC Conference on Personnel Assessment, St. Louis, MO.

Attachment D: Replication of the Basic Adverse Impact Analysis: Exam 7029 (1999)

Adverse Impact in Appointments by Race: Exam 7029 (1999)

This analysis replicates the first analysis above for the 1999 exam, but defines “appointed” differently and includes all test takers. Specifically, this analysis includes as appointed only those people with a disposition code of APP (i.e., appointed), which I take to mean actually on the job (since there is a separate disposition code, DEA, for those who declined appointment). This analysis also includes all people who took the exam, whether or not they dropped out of the post-exam screening process. I present only the two tables, without the introductory and interpretative narrative, since these are essentially the same as in the corresponding analysis in the body of my report.

Table D1. Appointments from Exam 7029, by Ethnic/Racial Group			
Group	Number Appointed (percent)	Number Not Appointed (percent)	Total (percent)
Black	100 (5.7%)	1,649 (94.3%)	1,749 (100%)
White	2,802 (21.7%)	10,113 (78.3%)	12,915 (100%)
Total	2,902	11,762	14,664

Table D2. Adverse Impact: Appointments from Exam 7029						
Black: % Appointed	White: % Appointed	Adverse Impact Ratio	Probability	Chi Square	z-score units	Shortfall
5.7%	21.7%	.26	< .0001	246.7	15.7	246

Attachment E: Replication of the Basic Adverse Impact Analysis: Exam 2043 (2002)

Adverse Impact in Appointments by Race: Exam 2043 (2002)

This analysis replicates the first analysis above for the 1999 exam, but defines "appointed" differently and includes all test takers. Specifically, this analysis includes as appointed only those people with a disposition code of APP (i.e., appointed), which I take to mean actually on the job (since there is a separate disposition code, DEA, for those who declined appointment). This analysis also includes all people who took the exam, whether or not they dropped out of the post-exam screening process. I present only the two tables, without the introductory and interpretative narrative, since these are essentially the same as in the corresponding analysis in the body of my report.

Table E1. Appointments from Exam 2043, by Racial Group			
Group	Number Appointed (percent)	Number Not Appointed (percent)	Total (percent)
Black	79 (5.7%)	1,314 (94.3%)	1,393 (100%)
White	1,822 (13.1%)	12,056 (86.9%)	13,878 (100%)
Total	1,901	13,370	15,271

Table E2. Adverse Impact: Appointments from Exam 2043					
Black: % Appointed	White: % Appointed	Adverse Impact Ratio	Probability	Chi Square	Standard Deviation Units
5.7%	13.1%	.44	< .0001	63.9	8.0